



UNIVERSITAT DE
BARCELONA

El rol de la sèrie documental en entorns de gestió dadificats: una proposta metodològica en el marc de la governança de la informació

Jordi Serra Serra

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (www.tdx.cat) i a través del Dipòsit Digital de la UB (diposit.ub.edu) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tdx.cat) y a través del Repositorio Digital de la UB (diposit.ub.edu) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (www.tdx.cat) service and by the UB Digital Repository (diposit.ub.edu) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.



UNIVERSITAT DE
BARCELONA

Tesi doctoral:

**El rol de la sèrie documental en
entorns de gestió d'adificats: una
proposta metodològica en el marc
de la governança de la informació**

Jordi Serra Serra

Director: Dr. Miquel Térmens Graells

Programa de Doctorat d'Informació i Comunicació

Departament de Biblioteconomia, Documentació i Comunicació
Audiovisual de la Universitat de Barcelona

Departament de Ciències de la Documentació i Història de la
Ciència de la Universitat de Saragossa

Barcelona, 2023

A la Sílvia i l'Anna

Als meus pares

Agraïments

En primer lloc, vull donar les gràcies a les persones que han fet possible des del Consorci d'Educació de Barcelona que les aportacions conceptuals d'aquesta tesi siguin una realitat aplicada i mesurable. Vull agrair molt especialment a Mercè Massa, Gerent del Consorci, i a Montse Simón, Directora d'Atenció a la Comunitat Educativa, la confiança dipositada en aquest projecte, i a tot l'equip de brillants professionals de la Unitat de Sistemes d'Informació i Disseny de Processos i de l'Àrea TIC funcional del Consorci la inestimable implicació en el desenvolupament del projecte.

Vull agrair també a Carles Drudis i al seu equip professional el suport tecnològic en el desenvolupament de la segona versió de l'eina de recollida, i l'ajuda en alguns aspectes del tractament de les dades.

El rerefons conceptual de la metodologia utilitzada en aquesta tesi es remunta a l'experiència de l'Arxiu Central del Departament de Cultura, i especialment als projectes portats a terme a l'extint Departament d'Universitats, Recerca i Societat de la Informació de la Generalitat de Catalunya (DURSI). Vull agrair doncs a Montserrat Canela, Isabel Campos i Joan Domingo, de l'equip de Cultura, i també a Montserrat Casellas, de l'Arxiu Central del DURSI, la possibilitat d'haver pogut impulsar plegats aquests projectes.

Aquesta tesi no hagués estat possible tampoc sense un context professional dinàmic i agosarat, com ha estat el de l'arxivística catalana en aquests darrers decennis. Sense les ganes d'innovar de tot un col·lectiu, cap evolució conceptual és possible. Vull donar les gràcies a molta gent de l'Associació d'Arxivers i Gestors de Documents de Catalunya amb qui he pogut col·laborar, i especialment a alguns professionals a qui admiro, i amb qui he tingut l'oportunitat de discutir aspectes clau d'aquesta tesi: a Lluís-Esteve Casellas per la concepció relacional de les entitats documentals, a Julio Quílez pels plantejaments ontològics, i a Antoni de Pàdua Puig-Pey, Pere Guiu i Joan Carles Faus, per la cabdal aportació de conceptes imbricats en la metodologia APS.

Vull agrair també a professionals d'altres països haver pogut discutir amb ells sobre alguns aspectes conceptuals i metodològics d'aquesta tesi, especialment amb Ulrich Kampffmeyer sobre determinats aspectes de la governança, amb Geoffrey Yeo sobre el concepte de document, i amb María Teresa Bermúdez i Raquel Umaña sobre l'aplicabilitat de les tècniques de mapatge exposades en les dues guies metodològiques.

L'experiència del Màster de Gestió Documental i Informació a les Empreses (MGDIE) a la Facultat de Biblioteconomia i Documentació (actualment Informació i Mitjans Audiovisuals) de la Universitat de Barcelona fou un veritable laboratori d'idees que va alimentar nombroses aportacions d'aquesta tesi. Vull agrair sobretot a Jordi Andreu, Núria Palomar, Pilar Campos, i a l'alumnat del Màster, les hores passades compartint visions sobre la gestió documental i, en la mesura del possible, fent-la evolucionar.

I finalment vull donar les gràcies a la persona que em va fer creure que podia fer aquesta tesi, Montserrat Sebastià, a qui em va tutoritzar el DEA, Ernest Abadal, i a qui m'ha acompanyat com a director en el desenvolupament de la tesi, ajudant-me en els moments difícils i orientant-me en tot moment amb els consells més encertats, Miquel Térmens. A tots ells, moltes gràcies.

Sumari

Capítol 1. Introducció.....	39
1.1 Context.....	39
1.1.1 La relació de la gestió documental amb la informació estructurada fins l'explosió del Big Data	39
1.1.2 Aparició del Big Data i els nous models de gestió de la informació estructurada.....	43
1.1.3 Conseqüències sobre la gestió documental i l'arxivística	46
1.2 Justificació de l'objecte i interès per a la disciplina arxivística.....	48
1.3 Hipòtesis i objectius.....	50
1.3.1 Objectius teòrics	50
1.3.1.1 Objectiu 1: característiques de la sèrie documental dadificada.....	50
1.3.1.2 Objectiu 2: característiques de l'ecosistema tecnològic de la sèrie documental dadificada	51
1.3.1.3 Objectiu 3: equivalència entre la sèrie documental i el sistema d'informació.....	51
1.3.1.4 Objectiu 4: les arquitectures distribuïdes com a entorns de gestió (transaccionals).....	51
1.3.1.5 Objectiu 5: equivalència entre tipus de sèries documentals i tipus de sistemes d'informació.....	52
1.3.1.6 Objectiu 6: les visualitzacions de dades com a recurs per a la gestió ...	52
1.3.1.7 Objectiu 7: de la classificació funcional a la classificació ontològica	53
1.3.2 Objectius instrumentals.....	53
1.4 Límits i abast.....	55
1.5 Metodologia.....	57
1.5.1 Fase 1: Anàlisi de l'estat de la qüestió	57
1.5.2 Fase 2: Preparació i recollida d'informació.....	58
1.5.2.1 Fixació dels criteris metodològics d'identificació	58
1.5.2.2 Desenvolupament del sistema de recollida i anàlisi.....	59
1.5.2.3 Tria del cas d'aplicació	60
1.5.3 Fase 3. Identificació, caracterització i mapatge del substrat informacional..	62
1.5.4 Fase 4. Anàlisi i transformació del substrat informacional	64

1.5.5 Fase 5. Anàlisi de les dades i extracció de conclusions	65
1.5.5.1 Anàlisi quantitativa per als objectius teòrics.....	65
1.5.5.2 Anàlisi qualitativa per als objectius instrumentals	66
1.6 Estructura de la tesi.....	67
1.6.1 Estructura dels capítols.....	67
1.6.2 Versió per a publicació.....	68
1.7 Guia de lectura.....	70
Capítol 2. Estat de la qüestió	71
2.1 Plantejament i estructura de l'estat de la qüestió.....	71
2.2 El constructe tecnològic: documents, dades i sistemes d'informació	75
2.2.1 La dialèctica document-dada i la crisi del concepte de document	75
2.2.2 Concepte i tipologia de la dada.....	82
2.2.2.1 Tipus de dades segons la informàtica i la ciència de dades.....	83
2.2.2.1.1 Dades transaccionals.....	84
2.2.2.1.2 Dades mestres.....	84
2.2.2.1.3 Dades referencials	88
2.2.2.1.4 Diferències entre els tipus de dades	88
2.2.2.2 Qualitat i valor de les dades: el valor evidencial de les dades	89
2.2.2.3 La descripció de les dades: els catàlegs de dades	93
2.2.3 Els sistemes d'informació.....	98
2.2.3.1 El model relacional com a referent dels sistemes de gestió.....	99
2.2.3.2 Els sistemes de gestió de dades mestres (MDM).....	100
2.2.3.3 Els models orientats a l'explotació: el modelat dimensional i el <i>business Intelligence</i>	102
2.2.3.4 Els models distribuïts: el model llac de dades i el <i>Big Data</i>	103
2.2.3.5 L'evolució del <i>data lake</i> : el <i>data fabric</i> i la virtualització	106
2.2.3.6 Els models orientats a objectes: present i futur dels gestors documentals	109
2.3 El constructe documental: la sèrie documental com a clau de volta de l'aportació arxivística	115
2.3.1 Característiques i evolució del concepte de sèrie documental a l'arxivística moderna	115
2.3.1.1 El concepte de sèrie documental i la seva estructura	117

2.3.1.2 La visió utilitarista de la sèrie documental: la sèrie com a recurs per a la gestió	119
2.3.1.3 El caràcter interseccional de la sèrie documental	121
2.3.1.4 Tipus de sèries documentals	124
2.3.1.4.1 Segons el tipus d'expedient que genera.....	124
2.3.1.4.2 Segons les relacions que mantenen entre elles	128
2.3.1.4.2.1 Relació estructural (funcional)	128
2.3.1.4.2.2 Relació de continuïtat (anterior/conseqüent)	129
2.3.1.4.2.3 Relació de derivació (primària/secundària)	129
2.3.1.4.2.4 Relació de dependència	129
2.3.1.4.2.5 Relació de contingut.....	130
2.3.2 El requisit d'existència de la sèrie documental: el context funcional	130
2.3.2.1 La classificació per funcions i activitats.....	132
2.3.2.2 El concepte de procés i la identificació de sèries documentals	134
2.3.2.3 La jerarquització dels processos i l'impacte sobre la classificació funcional.....	137
2.3.3 La crisi del model funcional.....	138
2.3.3.1 Crítica de la classificació jeràrquica i única.....	139
2.3.3.2 Crítica del criteri funcional en la classificació arxivística	141
2.3.4 L'utilitarisme extrem i els intents de simplificació: els <i>Big Buckets</i>	143
2.3.5 La introducció dels models ontològics en la classificació i la descripció arxivística.....	147
2.3.6 La dadificació de la sèrie documental	148
Capítol 3. Metodologia per a la construcció del sistema de governança	153
3.1 Introducció.....	153
3.1.1 Objectius.....	153
3.1.2 Bases metodològiques i models de referència	155
3.1.2.1 Models per a la construcció de sistemes de governança de la informació	155
3.1.2.2 Models per al mapatge de processos i de sèries documentals	157
3.1.2.3 Models per al mapatge de sistemes d'informació i aplicacions	158
3.1.2.4 La dadificació de processos i el mapatge de dades.....	159
3.1.3 Llicències terminològiques	160

3.2 Guia d'inventari i mapatge (GMAP)	163
3.2.1 Estructura i mètode d'aquesta guia	163
3.2.2 Model conceptual per a la construcció dels mapes	164
3.2.3 Pas 1: mapatge dels agents.....	168
3.2.3.1 Què s'entén per agent	168
3.2.3.2 Com s'identifica un agent	169
3.2.3.3 Com es caracteritzen els agents.....	170
3.2.3.4 Com es relacionen els agents entre ells	170
3.2.3.5 Com es relacionen els agents amb altres entitats.....	171
3.2.3.6 Com es gestionen les altes, baixes i modificacions d'agents	172
3.2.4 Pas 2: mapatge dels processos	174
3.2.4.1 Què s'entén per procés	174
3.2.4.2 Com s'identifica un procés.....	174
3.2.4.2.1 Buidatge de funcions i activitats (processos).....	176
3.2.4.2.2 Identificació i eliminació de duplicitats	179
3.2.4.2.3 Establiment de relacions	180
3.2.4.2.4 Validació mitjançant iteracions	183
3.2.4.2.5 Codificació i denominació definitiva.....	184
3.2.4.3 Com es caracteritzen els processos	184
3.2.4.3.1 Caracterització intrínseca.....	185
3.2.4.3.2 Caracterització per descomposició.....	186
3.2.4.4 Com es relacionen els processos entre ells.....	186
3.2.4.5 Com es relacionen els processos amb altres entitats	188
3.2.4.6 Com es gestionen les altes, baixes i modificacions de processos	189
3.2.5 Pas 3: mapatge dels sistemes d'informació	191
3.2.5.1 Què s'entén per sistema d'informació.....	191
3.2.5.2 Com s'identifica un sistema d'informació	191
3.2.5.2.1 Determinació de l'abast: quins sistemes s'inclouen.....	192
3.2.5.2.2 Determinació dels límits: components del sistema	193
3.2.5.3 Com es caracteritzen els sistemes	194
3.2.5.4 Com es relacionen els sistemes entre ells	196
3.2.5.5 Com es relacionen els sistemes d'informació amb altres entitats	198

3.2.5.6 Com es gestionen les altes, baixes i modificacions de sistemes d'informació.....	199
3.2.6 Pas 4: mapatge de les sèries documentals.....	201
3.2.6.1 Què s'entén per sèrie documental.....	201
3.2.6.2 Com s'identifica una sèrie documental.....	203
3.2.6.2.1.1 Identificació de la sèrie primària teòrica.....	203
3.2.6.2.1.2 Identificació de les sèries reals.....	203
3.2.6.3 Com es caracteritzen les sèries documentals.....	204
3.2.6.3.1 Caracterització intrínseca.....	205
3.2.6.3.1.1 Tipificació de la sèrie.....	205
3.2.6.3.1.2 Assignació del criteri de nominalització.....	206
3.2.6.3.1.3 Determinació de la vitalitat o de la sèrie.....	207
3.2.6.3.1.4 Determinació del cicle de vida de la sèrie.....	208
3.2.6.3.1.5 Determinació del règim d'accés de la sèrie.....	209
3.2.6.3.2 Caracterització per descomposició.....	209
3.2.6.4 Com es relacionen les sèries entre elles.....	210
3.2.6.5 Com es relacionen les sèries documentals amb altres entitats.....	212
3.2.6.6 Com es gestionen les altes, baixes i modificacions de sèries documentals.....	212
3.2.7 Pas 5: mapatge de dades.....	215
3.2.7.1 Què s'entén per dada.....	215
3.2.7.2 Com s'identifica una dada.....	216
3.2.7.2.1 Identificar els objectes descrits.....	216
3.2.7.2.2 Identificar les dades elementals.....	218
3.2.7.2.3 Harmonitzar les dades.....	218
3.2.7.3 Com es caracteritzen les dades.....	219
3.2.7.4 Com es relacionen les dades entre elles.....	220
3.2.7.5 Com es relacionen les dades amb altres entitats.....	220
3.2.7.6 Com es gestionen les altes, baixes i modificacions de dades.....	221
3.3 Guia de dadificació de processos (GDAD).....	223
3.3.1 Estructura i mètode d'aquesta guia.....	223
3.3.2 Pas 1: descripció.....	226
3.3.2.1 Descriure i, si escau, diagramar el procés.....	226

3.3.2.1.1 Identificació de les tasques	227
3.3.2.1.2 Descripció de les tasques	229
3.3.2.2 Derivar el circuit documental del procés	232
3.3.2.2.1 Identificació de les evidències	233
3.3.2.2.2 Descripció de les evidències	234
3.3.2.3 Identificar les tasques substantives del procés	236
3.3.2.4 Agrupar en sèries documentals les evidències del procés	237
3.3.2.5 Identificar els agents promotors o competents del procés	238
3.3.2.6 Caracteritzar i relacionar les sèries documentals del procés	238
3.3.3 Pas 2: optimització	240
3.3.3.1 Analitzar les evidències essencials de les sèries primàries	241
3.3.3.2 Modificar el procés per a que generi el resultat documental òptim.....	241
3.3.4 Pas 3: deshibridació.....	243
3.3.4.1 Identificar les evidències en formats no digitals i transformar-ne el format	243
3.3.5 Pas 4: automatització.....	244
3.3.5.1 Incorporar un gestor documental com a destí de totes les evidències essencials	244
3.3.5.2 Valorar el tractament de la missatgeria.....	244
3.3.5.3 Suprimir les redundàncies documentals en les evidències essencials	245
3.3.5.4 Modificar el procés per a que generi un circuit d'expedient únic	246
3.3.5.5 Implementar el procediment en un gestor documental	247
3.3.6 Pas 5: dadificació.....	249
3.3.6.1 Incorporar un sistema d'informació com a destí de totes les evidències essencials	249
3.3.6.2 Construir el model (lògic) de dades del sistema d'informació que donarà suport al procés.....	250
3.3.6.2.1 Identificació dels objectes descrits	252
3.3.6.2.2 Extracció de les variables de cada evidència	253
3.3.6.2.3 Construcció de la matriu de variables del procés	258
3.3.6.2.4 Mapatge amb la matriu de variables global de l'organització.....	260
3.3.6.2.5 Modificació de les variables d'entrada del procés	261
3.3.6.2.6 Afegint valor al procés.....	262

3.3.6.3	Modificar el procés per a que generi el circuit documental dadificat ...	263
3.3.6.4	Implementar el procediment en un sistema d'informació	264
3.4	Annex: proposta de sistematització de la recollida d'informació	265
3.4.1	Anàlisi prèvia	265
3.4.2	Inventari i prioritització de fonts d'informació.....	266
3.4.2.1	Fonts personals.....	268
3.4.2.1.1	Entrevistes formals	268
3.4.2.1.1.1	Determinar/decidir el tipus d'entrevista i a qui entrevistar	268
3.4.2.1.1.2	Programar les entrevistes.....	269
3.4.2.1.1.3	Preparar les preguntes	269
3.4.2.1.1.4	Conduir les entrevistes.....	270
3.4.2.1.1.5	Recollir els resultats de les entrevistes.....	270
3.4.2.1.2	Qüestionaris.....	271
3.4.2.1.3	Entrevistes informals.....	271
3.4.2.1.4	Observació directa	272
3.4.2.2	Fonts documentals	272
3.4.2.2.1	Documents públics.....	272
3.4.2.2.2	Documents sobre constitució i planificació	273
3.4.2.2.3	Documents de recapitulació i memòria	273
3.4.2.2.4	Normativa i procediment	273
3.4.2.2.5	Projectes de millora.....	274
3.4.2.2.6	Formularis, models de documents i altres pautes formals.....	274
3.4.2.2.7	Sistemes d'informació i models de dades	275
3.4.2.2.8	Documentació en general	275
Capítol 4.	Eina de recollida i model de dades	277
4.1	Introducció.....	277
4.2	Versions	277
4.3	Funcionalitats	279
4.3.1	Funcionalitats d'administració	279
4.3.2	Funcionalitats de manteniment de dades.....	279
4.3.3	Funcionalitats d'interoperabilitat del sistema	283
4.3.4	Funcionalitats d'explotació del sistema	285

4.4 Model de dades.....	287
4.5 Models de referència i abreviatures per citació.....	291
4.6 Descripció dels components.....	300
4.6.1 Entitats.....	301
4.6.1.1 Entitat: Agent.....	301
4.6.1.1.1 Descripció de l'entitat.....	301
4.6.1.1.2 Atribut: Identificador.....	302
4.6.1.1.3 Atribut: Nom.....	303
4.6.1.1.4 Atribut: Tipus.....	304
4.6.1.1.5 Atribut: Contacte.....	305
4.6.1.1.6 Atribut: Observacions.....	305
4.6.1.1.7 Atribut: Data creació.....	306
4.6.1.1.8 Atribut: Data supressió.....	307
4.6.1.2 Entitat: Procés.....	309
4.6.1.2.1 Descripció de l'entitat.....	309
4.6.1.2.2 Atribut: Identificador.....	310
4.6.1.2.3 Atribut: Nom.....	311
4.6.1.2.4 Atribut: Descripció.....	312
4.6.1.2.5 Atribut: Tipus.....	313
4.6.1.2.6 Atribut: Família.....	313
4.6.1.2.7 Atribut: Àmbit.....	313
4.6.1.2.8 Atribut: Tipus iniciació.....	314
4.6.1.2.9 Atribut: Tipus finalització.....	314
4.6.1.2.10 Atribut: Tipus recurs.....	315
4.6.1.2.11 Atribut: Requisits a acreditar.....	316
4.6.1.2.12 Atribut: Terminis rellevants.....	316
4.6.1.2.13 Atribut: Necessitat de pagament.....	317
4.6.1.2.14 Atribut: Data creació.....	317
4.6.1.2.15 Atribut: Data supressió.....	318
4.6.1.3 Entitat: Tasca.....	320
4.6.1.3.1 Descripció de l'entitat.....	320
4.6.1.3.2 Atribut: Identificador.....	321
4.6.1.3.3 Atribut: Nom.....	321

4.6.1.3.4 Atribut: Descripció.....	322
4.6.1.3.5 Atribut: Tipus executiu.....	322
4.6.1.3.6 Atribut: Tipus en el flux.....	323
4.6.1.3.7 Atribut: Condició de bifurcació.....	324
4.6.1.3.8 Atribut: Data creació.....	324
4.6.1.3.9 Atribut: Data supressió.....	325
4.6.1.4 Entitat: Evidència.....	326
4.6.1.4.1 Descripció de l'entitat.....	326
4.6.1.4.2 Atribut: Identificador.....	328
4.6.1.4.3 Atribut: Nom.....	329
4.6.1.4.4 Atribut: Tipus.....	330
4.6.1.4.5 Atribut: Caràcter.....	330
4.6.1.4.6 Atribut: Format.....	331
4.6.1.4.7 Atribut: Autenticació.....	333
4.6.1.4.8 Atribut: Registre d'entrada.....	334
4.6.1.4.9 Atribut: Registre de sortida.....	334
4.6.1.4.10 Atribut: Essencial.....	335
4.6.1.4.11 Atribut: Data creació.....	336
4.6.1.4.12 Atribut: Data supressió.....	336
4.6.1.5 Entitat: Model de document.....	338
4.6.1.5.1 Descripció de l'entitat.....	338
4.6.1.5.2 Atribut: Identificador.....	338
4.6.1.5.3 Atribut: Versió.....	339
4.6.1.5.4 Atribut: Nom.....	339
4.6.1.5.5 Atribut: Tipus.....	340
4.6.1.5.6 Atribut: Format.....	341
4.6.1.5.7 Atribut: Data creació.....	342
4.6.1.5.8 Atribut: Data supressió.....	342
4.6.1.6 Entitat: Dada elemental.....	344
4.6.1.6.1 Descripció de l'entitat.....	344
4.6.1.6.2 Atribut: Identificador.....	345
4.6.1.6.3 Atribut: Nom.....	345
4.6.1.6.4 Atribut: Tipus de dada.....	346

4.6.1.6.5 Atribut: Descripció	346
4.6.1.6.6 Atribut: Regles de validació	346
4.6.1.6.7 Atribut: Nivell de seguretat RGPD	347
4.6.1.6.8 Atribut: Disponible per interoperabilitat externa	348
4.6.1.7 Entitat: Sèrie documental.....	349
4.6.1.7.1 Descripció de l'entitat	349
4.6.1.7.2 Atribut: Identificador	351
4.6.1.7.3 Atribut: Nom	352
4.6.1.7.4 Atribut: Criteri de nominalització	353
4.6.1.7.5 Atribut: Caràcter	354
4.6.1.7.6 Atribut: Descripció	355
4.6.1.7.7 Atribut: Vitalitat.....	356
4.6.1.7.8 Atribut: 100% Digital.....	356
4.6.1.7.9 Atribut: Data creació.....	357
4.6.1.7.10 Atribut: Data supressió	358
4.6.1.8 Entitat: Regla d'accés.....	360
4.6.1.8.1 Descripció de l'entitat	360
4.6.1.8.2 Atribut: Identificador	361
4.6.1.8.3 Atribut: Esdeveniment d'inici o disparador.....	361
4.6.1.8.4 Atribut: Termini (o període)	362
4.6.1.8.5 Atribut: Resolució (o disposició)	362
4.6.1.8.6 Atribut: Motiu de limitació	363
4.6.1.8.7 Atribut: Criteri d'aplicació de la regla	364
4.6.1.8.8 Atribut: Criteri de publicitat per transparència.....	364
4.6.1.8.9 Atribut: Data creació.....	365
4.6.1.8.10 Atribut: Data supressió	365
4.6.1.9 Entitat: Regla de conservació	366
4.6.1.9.1 Descripció de l'entitat	366
4.6.1.9.2 Atribut: Identificador	367
4.6.1.9.3 Atribut: Esdeveniment d'inici o disparador.....	367
4.6.1.9.4 Atribut: Termini (o període)	368
4.6.1.9.5 Atribut: Resolució (o disposició)	368
4.6.1.9.6 Atribut: Criteri de mostreig.....	369

4.6.1.9.7 Atribut: Suport de destí	370
4.6.1.9.8 Atribut: Format de destí.....	370
4.6.1.9.9 Atribut: Lloc de destí	371
4.6.1.9.10 Atribut: Data creació.....	371
4.6.1.9.11 Atribut: Data supressió	372
4.6.1.10 Entitat: Disposició normativa.....	373
4.6.1.10.1 Descripció de l'entitat	373
4.6.1.10.2 Atribut: Identificador	373
4.6.1.10.3 Atribut: Títol.....	374
4.6.1.10.4 Atribut: Rang	375
4.6.1.10.5 Atribut: Àmbit	375
4.6.1.10.6 Atribut: Data vigència	376
4.6.1.10.7 Atribut: Data derogació	376
4.6.1.11 Entitat: Sistema d'informació o aplicació.....	378
4.6.1.11.1 Descripció de l'entitat	378
4.6.1.11.2 Atribut: Identificador (codi)	380
4.6.1.11.3 Atribut: Nom	380
4.6.1.11.4 Atribut: Nom curt	381
4.6.1.11.5 Atribut: Responsable funcional.....	381
4.6.1.11.6 Atribut: Gestor TIC	382
4.6.1.11.7 Atribut: Proveïdor	382
4.6.1.11.8 Atribut: Entorn	383
4.6.1.11.9 Atribut: Entorn tecnològic	383
4.6.1.11.10 Atribut: Model de servei.....	384
4.6.1.11.11 Atribut: Abast	385
4.6.1.11.12 Atribut: Restriccions semàntiques	386
4.6.1.11.13 Atribut: Restriccions de format	386
4.6.1.11.14 Atribut: Controls d'integritat	386
4.6.1.11.15 Atribut: Controls d'accés	387
4.6.1.11.16 Atribut: Classificació de seguretat	388
4.6.1.11.17 Atribut: Descripció	388
4.6.1.11.18 Atribut: Funcionalitats.....	389
4.6.1.11.19 Atribut: Ubicació.....	390

4.6.1.11.20 Atribut: Data inici servei.....	390
4.6.1.11.21 Atribut: Data fi servei.....	391
4.6.1.12 Entitat: Interfície de connexió entre sistemes.....	392
4.6.1.12.1 Descripció de l'entitat.....	392
4.6.1.12.2 Atribut: Identificador (codi).....	393
4.6.1.12.3 Atribut: Nom.....	393
4.6.1.12.4 Atribut: Tipus d'interfície.....	394
4.6.1.12.5 Atribut: Tipus d'importació.....	394
4.6.1.12.6 Atribut: Procediment d'importació.....	395
4.6.1.12.7 Atribut: Actualització.....	395
4.6.1.12.8 Atribut: Historificació.....	396
4.6.1.12.9 Atribut: Nom del fitxer de càrrega.....	397
4.6.1.12.10 Atribut: Restriccions semàntiques (dades).....	397
4.6.1.12.11 Atribut: Restriccions de format.....	398
4.6.1.12.12 Atribut: Observacions.....	398
4.6.1.12.13 Atribut: Data inici servei.....	398
4.6.1.12.14 Atribut: Data fi servei.....	399
4.6.1.13 Entitat: Visualització de dades.....	400
4.6.1.13.1 Descripció de l'entitat.....	400
4.6.1.13.2 Atribut: Identificador (codi).....	400
4.6.1.13.3 Atribut: Nom.....	401
4.6.1.13.4 Atribut: Tipus.....	401
4.6.1.13.5 Atribut: Finalitat.....	401
4.6.1.13.6 Atribut: Destinataris.....	402
4.6.1.13.7 Atribut: Usuaris.....	402
4.6.1.13.8 Atribut: Data inici servei.....	403
4.6.1.13.9 Atribut: Data fi servei.....	403
4.6.1.14 Entitat: Objecte.....	405
4.6.1.14.1 Descripció de l'entitat.....	405
4.6.1.14.2 Atribut: Nom.....	406
4.6.1.14.3 Atribut: Tipus.....	406
4.6.2 Relacions.....	407
4.6.2.1 Relació: agent amb agent.....	407

4.6.2.1.1	Descripció de la relació	407
4.6.2.1.2	Atribut: Rol de l'agent.....	407
4.6.2.2	Relació: agent amb procés	408
4.6.2.2.1	Descripció de la relació	408
4.6.2.2.2	Atribut: Rol de l'agent.....	408
4.6.2.3	Relació: agent amb tasca	409
4.6.2.3.1	Descripció de la relació	409
4.6.2.4	Relació: agent amb sèrie documental.....	410
4.6.2.4.1	Descripció de la relació	410
4.6.2.5	Relació: agent amb regla d'accés.....	410
4.6.2.5.1	Descripció de la relació	410
4.6.2.6	Relació: agent amb sistema d'informació.....	411
4.6.2.6.1	Descripció de la relació	411
4.6.2.7	Relació: agent amb evidència.....	412
4.6.2.7.1	Descripció de la relació	412
4.6.2.7.2	Atribut: Rol de l'agent.....	412
4.6.2.8	Relació: procés amb procés	413
4.6.2.8.1	Descripció de la relació	413
4.6.2.8.2	Atribut: Relació de continuïtat	414
4.6.2.8.3	Atribut: Relació de comprensió	414
4.6.2.9	Relació: procés amb tasca	415
4.6.2.9.1	Descripció de la relació	415
4.6.2.10	Relació: procés amb disposició normativa	416
4.6.2.10.1	Descripció de la relació	416
4.6.2.11	Relació: tasca amb tasca.....	416
4.6.2.11.1	Descripció de la relació	416
4.6.2.11.2	Atribut: Rol de la tasca	417
4.6.2.12	Relació: sèrie documental amb sèrie documental.....	417
4.6.2.12.1	Descripció de l'entitat	417
4.6.2.12.2	Atribut: Rol de la sèrie	418
4.6.2.13	Relació: procés amb sèrie documental.....	419
4.6.2.13.1	Descripció de la relació	419
4.6.2.14	Relació: sèrie documental amb regla d'accés.....	420

4.6.2.14.1	Descripció de la relació	420
4.6.2.15	Relació: sèrie documental amb regla de conservació	421
4.6.2.15.1	Descripció de la relació	421
4.6.2.16	Relació: procés amb evidència	421
4.6.2.16.1	Descripció de la relació	421
4.6.2.17	Relació: tasca amb evidència	422
4.6.2.17.1	Descripció de la relació	422
4.6.2.17.2	Atribut: Sentit de l'evidència	422
4.6.2.18	Relació: sèrie documental amb evidència	423
4.6.2.18.1	Descripció de la relació	423
4.6.2.19	Relació: sistema d'informació amb evidència	424
4.6.2.19.1	Descripció de la relació	424
4.6.2.20	Relació: evidència amb dada elemental	424
4.6.2.20.1	Descripció de l'entitat	424
4.6.2.20.2	Atribut: Tipus d'ocurrència	425
4.6.2.21	Relació: evidència amb model de document	426
4.6.2.21.1	Descripció de la relació	426
4.6.2.22	Relació: interfície amb dada elemental	426
4.6.2.22.1	Descripció de la relació	426
4.6.2.23	Relació: sistema d'informació amb dada elemental	427
4.6.2.23.1	Descripció de la relació	427
4.6.2.24	Relació: sistema d'informació amb objecte	427
4.6.2.24.1	Descripció de la relació	427
4.6.2.25	Relació: objecte amb objecte	428
4.6.2.25.1	Descripció de la relació	428
4.6.2.26	Relació: objecte amb dada elemental	429
4.6.2.26.1	Descripció de la relació	429
4.6.2.27	Relació: Visualització de dades amb dada elemental	429
4.6.2.27.1	Descripció de la relació	429
4.6.2.28	Relació: sistema d'informació amb sistema d'informació (interfície)..	430
4.6.3	Correspondència de les entitats amb cada estàndard	431
Capítol 5.	Proposta d'indicadors per a la governança de la informació	433
5.1	Plantejament metodològic	433

5.2	Indicadors sobre les sèries documentals	439
5.2.1	Mètriques descriptives de les sèries documentals.....	439
5.2.2	Mètriques derivades de la relació de la sèrie documental amb altres entitats	439
5.2.3	Mètriques sobre les dades de la sèrie documental.....	440
5.2.4	Indicadors d'anàlisi de les sèries documentals	441
5.2.4.1	Índex d'influència de la sèrie documental	441
5.2.4.2	Índex de dadificació de la sèrie documental	445
5.2.4.3	Índex de ròssec	447
5.2.4.3.1	Ròssec intrasèrie	448
5.2.4.3.2	Ròssec intersèrie	450
5.2.4.4	Índex de transaccionalitat de la sèrie documental	454
5.2.4.4.1	Índex de transaccionalitat calculat a partir de l'objecte de les dades	455
5.2.4.4.2	Índex de transaccionalitat calculat a partir del tipus de dades	455
5.2.4.4.3	Índex de transaccionalitat calculat a partir dels objectes	456
5.2.5	Mapa de relacions entre sèries documentals	456
5.3	Indicadors sobre els processos	457
5.3.1	Mètriques descriptives dels processos.....	457
5.3.2	Mètriques derivades de les relacions del procés amb altres entitats	457
5.3.3	Mètriques sobre les dades del procés.....	458
5.3.4	Indicadors d'anàlisi dels processos	459
5.3.4.1	Índex d'influència del procés	459
5.3.4.2	Índex de dadificació del procés.....	461
5.3.4.3	Índex de ròssec	463
5.3.4.3.1	Ròssec intraprocés	463
5.3.4.3.2	Ròssec interprocés	464
5.3.5	Mapa de les relacions entre processos	465
5.4	Indicadors sobre els sistemes d'informació	466
5.4.1	Mètriques descriptives dels sistemes d'informació	466
5.4.2	Mètriques de la transformació	466
5.4.3	Mètriques derivades de la relació del sistema d'informació amb altres entitats	467

Capítol 6. Cas d'aplicació: el sistema de governança de la informació del Consorci d'Educació de Barcelona	493
6.1 L'objecte d'anàlisi: el Consorci d'Educació de Barcelona	493
6.2 Aplicació de la metodologia i els indicadors al Consorci d'Educació de Barcelona	495
6.2.1 Objectius del cas d'aplicació	495
6.2.2 Base tecnològica del cas d'aplicació	495
6.3 Estat de la recollida de dades.....	496
6.4 Indicadors sobre les sèries documentals	504
6.4.1 Mètriques descriptives de les sèries documentals.....	504
6.4.2 Mètriques derivades de la relació de la sèrie documental amb altres entitats	509
6.4.3 Mètriques sobre les dades de la sèrie documental.....	512
6.4.4 Indicadors d'anàlisi de les sèries documentals	517
6.4.4.1 Índex d'influència de la sèrie documental	517
6.4.4.2 Índex de dadificació de la sèrie documental	518
6.4.4.3 Índex de ròssec	521
6.4.4.4 Índex de transaccionalitat	522
6.4.5 Mapa de relacions entre sèries documentals	525
6.5 Indicadors sobre els processos	526
6.5.1 Mètriques descriptives dels processos.....	526
6.5.2 Mètriques derivades de les relacions del procés amb altres entitats	527
6.5.3 Mètriques sobre les dades del procés.....	531
6.5.4 Indicadors d'anàlisi dels processos	533
6.5.4.1 Índex d'influència del procés	533
6.5.4.2 Índex de dadificació del procés.....	534
6.5.4.3 Índex de ròssec	534
6.5.5 Mapa de les relacions entre processos	535
6.6 Indicadors sobre els sistemes d'informació	536
6.6.1 Mètriques descriptives dels sistemes d'informació	536
6.6.2 Mètriques de la transformació.....	545
6.6.3 Mètriques derivades de la relació del sistema d'informació amb altres entitats	547

6.6.4	Mètriques sobre les dades del sistema d'informació	553
6.6.5	Indicadors d'anàlisi dels sistemes d'informació	557
6.6.5.1	Índex d'influència del sistema d'informació	557
6.6.5.1.1	Índex de connectivitat potencial	559
6.6.5.2	Índex de transaccionalitat del sistema	559
6.6.5.3	Classificació dels sistemes pel nivell de seguretat de les dades.....	561
6.6.6	Mapa de relacions entre sistemes d'informació.....	563
6.7	Indicadors sobre les dades.....	564
6.7.1	Mètriques descriptives de les dades elementals	564
6.7.2	Mètriques derivades de la relació de la dada elemental amb altres entitats	566
6.7.2.1.1	A nivell de dada elemental	566
6.7.2.1.2	A nivell d'objecte	567
6.7.2.1.3	A nivell de superobjecte	567
6.7.2.1.4	A nivell de tipus d'objecte (cosa/acció).....	568
6.7.3	Indicadors d'anàlisi de les dades elementals	569
6.7.3.1	Índex de rellevància de la dada elemental.....	569
6.7.3.2	Índex de rellevància de l'objecte i del superobjecte	570
6.7.3.2.1	Índex de rellevància global.....	570
6.7.4	Mapa de relacions entre dades i objectes	572
6.7.5	Mapa global de relacions entre entitats.....	575
Capítol 7.	Anàlisi dels resultats del cas d'aplicació	577
7.1	Anàlisi de les característiques de les sèries documentals dadificades del Consorci d'Educació de Barcelona.....	577
7.1.1	Sèrie documental ESC001-001 Seguiment de falsos empadronaments ...	582
7.1.2	Sèrie documental GRH023-001 Gestió de borses de personal de suport als centres educatius (COVID-19).....	584
7.1.3	Sèrie documental INF001-001 Explotació de dades.....	587
7.1.4	Sèrie documental INF009-001 Manteniment de les dades de la Base de Dades de Centres (BDC)	590
7.1.5	Sèrie documental LOP003-001 Registre i notificacions de violacions de seguretat	594
7.1.6	Sèrie documental LOP005-001 Auditories APDCAT i controls de seguretat sobre protecció de dades de caràcter personal	597

7.1.7 Sèrie documental NEE002-001 Actuacions en casos de sospita de maltractament i/o assetjament en un centre educatiu	599
7.1.8 Sèrie documental NEE021-001 Seguiment i valoració d'alumnes en situació de vulnerabilitat i/o malestar social o psicosocial	602
7.1.9 Sèrie documental ORG003-001 Anàlisi i millora de processos.....	606
7.1.10 Sèrie documental ORG004-001 Normalització de formularis i models de documents	608
7.2 Objectiu 1: característiques de la sèrie documental dadificada.....	612
7.3 Objectiu 2: característiques de l'ecosistema tecnològic de la sèrie documental dadificada.....	619
7.4 Objectiu 3: equivalència entre la sèrie documental i el sistema d'informació ...	621
7.4.1 Equivalència amb el model relacional	621
7.4.2 Equivalència amb el model <i>data warehouse</i>	622
7.4.3 Equivalència amb el model <i>data lake</i>	622
7.4.4 Equivalència amb el model orientat a objectes	623
7.4.5 Conclusió	624
7.5 Objectiu 4: les arquitectures distribuïdes com a entorns de gestió (transaccionals).....	625
7.5.1 Conclusió	637
7.6 Objectiu 5: equivalència entre tipus de sèries documentals i tipus de sistemes d'informació.....	638
7.6.1 Conclusió	645
7.7 Objectiu 6: les visualitzacions de dades com a recurs per a la gestió.....	646
7.7.1 Conclusió	652
7.8 Objectiu 7: de la classificació funcional a la classificació ontològica	653
7.8.1 Conclusió	658
7.9 Epíleg del cas d'aplicació	658
Capítol 8. Conclusions generals	659
8.1 Objectius teòrics.....	660
8.1.1 Objectiu 1: característiques de la sèrie documental dadificada	660
8.1.2 Objectiu 2: característiques de l'ecosistema tecnològic de la sèrie documental dadificada	665
8.1.2.1 Model tradicional	665
8.1.2.2 Model <i>data fabric</i>	666

8.1.2.3 Proposta d'indicadors per a decidir el model	667
8.1.2.4 Conseqüències per a la gestió documental	669
8.1.3 Objectiu 3: equivalència entre la sèrie documental i el sistema d'informació	671
8.1.4 Objectiu 4: les arquitectures distribuïdes com a entorns de gestió (transaccionals)	678
8.1.5 Objectiu 5: equivalència entre tipus de sèries documentals i tipus de sistemes d'informació	679
8.1.6 Objectiu 6: les visualitzacions de dades com a recurs per a la gestió	681
8.1.7 Objectiu 7: de la classificació funcional a la classificació ontològica	683
8.2 Objectius instrumentals	690
8.2.1 Punts febles	690
8.2.2 Punts forts	693
8.2.3 Recomanacions	694
8.2.4 Futures línies de recerca.....	695
8.3 Epíleg: resum de les conclusions generals.....	696
8.3.1 Objectiu 1: característiques de la sèrie documental dadificada	696
8.3.2 Objectiu 2: característiques de l'ecosistema tecnològic de la sèrie documental dadificada	696
8.3.3 Objectiu 3: equivalència entre la sèrie documental i el sistema d'informació	697
8.3.4 Objectiu 4: les arquitectures distribuïdes com a entorns de gestió (transaccionals)	697
8.3.5 Objectiu 5: equivalència entre tipus de sèries documentals i tipus de sistemes d'informació	698
8.3.6 Objectiu 6: les visualitzacions de dades com a recurs per a la gestió	698
8.3.7 Objectiu 7: de la classificació funcional a la classificació ontològica	698
Capítol 9. Bibliografia.....	701

Índex de figures¹

Figura 2-1: Estructura conceptual del model data fabric (Font: Delgado, 2022)	108
Figura 2-2: Jerarquia d'agrupacions documentals ISAD(G)	117
Figura 2-3: Perspectiva d'un tràmit només a partir de la jerarquia d'agrupacions documentals	123
Figura 2-4: Perspectiva d'un tràmit a partir de la intersecció entre la funció i l'òrgan	123
Figura 2-5: Representació de les possibles relacions entre sèries documentals	130
Figura 2-6: Esquema del model de procediment seqüencial	137
Figura 2-7: Esquema del model de procediment radial	137
Figura 2-8: Esquema dels dos models d'anàlisi de la norma ISO 26122	138
Figura 3-1: Estructura general de la metodologia	166
Figura 3-2: Estructura de la Guia d'inventari i mapatge (GMAP)	167
Figura 3-3: Estructura per al mapatge dels agents	168
Figura 3-4: Estructura per al mapatge dels processos	174
Figura 3-5: Patró de descomposició de l'anàlisi funcional	176
Figura 3-6: Nivells inicials per a la identificació de processos	177
Figura 3-7: Estructura per al buidatge dels processos	177
Figura 3-8: Detecció de duplicitats en l'anàlisi funcional	180
Figura 3-9: Compactació del model funcional	182
Figura 3-10: Establiment de relacions entre processos	182
Figura 3-11: Estructura per al mapatge dels sistemes d'informació	191
Figura 3-12: Estructura per al mapatge de les sèries documentals	201
Figura 3-13: Estructura per al mapatge de les dades	215
Figura 3-14: Estructura de la Guia de dadificació de processos (GDAD)	225
Figura 3-15: Estructura per a la descripció i diagramació del procés	226
Figura 3-16: Estructura per a la derivació del circuit documental del procés	232
Figura 3-17: Pautes d'actuació per optimitzar el resultat documental del procés	242
Figura 3-18: Pautes d'actuació per transformar el format de les evidències a digital.	243
Figura 3-19: Estructura per a la dadificació del procés	249
Figura 4-1: Menú de gestió de processos	280
Figura 4-2: Menú de gestió de sistemes d'informació	280
Figura 4-3: Menú de gestió d'interaccions entre sistemes d'informació	281
Figura 4-4: Il·lustració 4. Menú de gestió de les entitats documentals	282
Figura 4-5: Exemple de fitxa de manteniment de dades	282
Figura 4-6: Opcions d'exportació de les llistes	283
Figura 4-7: Exemple de fitxa descriptiva de l'entitat Procés	283
Figura 4-8: Exemple de guia de procediment	284
Figura 4-9: Exemple d'exportació de fluxograma de procediment	285

¹ Totes les figures, gràfics i taules d'aquesta tesi, llevat d'un cas, han estat elaborades per l'autor. Per aquest motiu, només en aquest cas se cita la font (figura 2-1).

Figura 4-10: Exemple d'opcions de cerca i filtrat d'informació	286
Figura 4-11: Exemple d'entorn de visualització: mapa global de processos, sèries documentals i sistemes d'informació.....	286
Figura 4-12: Diagrama entitat-relació del model de dades de l'eina CATPROC (model lògic).....	300
Figura 5-1: Exemple del mapa de processos del Consorci d'Educació de Barcelona	437
Figura 5-2: Exemple del mapa global d'actius d'informació del Consorci d'Educació de Barcelona	437
Figura 5-3: Mapa de relacions de la sèrie documental ATU004-003	443
Figura 5-4: Mapa de relacions del procés ATU004	444
Figura 5-5: Mapa de relacions de la sèrie documental GRH016-001	444
Figura 5-6: Mapa de relacions del procés GRH016	444
Figura 5-7: Relacions entre tres sèries documentals del procés INF001	451
Figura 5-8: Exemple d'anàlisi visual de coincidències en dades elementals entre sèries documentals i sistemes d'informació.....	454
Figura 5-9: Mapa de dades elementals, evidències i sistemes d'informació de l'objecte "Assessorament a un centre educatiu en matèria de gestió documental"	487
Figura 5-10: Mapa de dades elementals, evidències i sistemes d'informació de l'objecte "Auditoria de protecció de dades personals"	488
Figura 6-1: Mapa d'objectes de la sèrie documental ORG002-001	516
Figura 6-2: Mapa d'objectes de la sèrie documental INF001-003.....	516
Figura 6-3: Fragment del mapa de sèries documentals del Consorci d'Educació de Barcelona	525
Figura 6-4: Fragment del mapa de processos del Consorci d'Educació de Barcelona	535
Figura 6-5: Mapa d'interaccions entre sistemes d'informació del Consorci d'Educació de Barcelona	557
Figura 6-6: Fragment del mapa de sistemes d'informació del Consorci d'Educació de Barcelona	563
Figura 6-7: Superobjectes descrits per un major nombre de dades elementals	564
Figura 6-8: Fragment del mapa de dades elementals del Consorci d'Educació de Barcelona	572
Figura 6-9: Fragment del mapa d'objectes del Consorci d'Educació de Barcelona ...	573
Figura 6-10: Fragment del mapa d'interaccions entre objectes del Consorci d'Educació de Barcelona	574
Figura 6-11: Fragment del mapa global d'actius d'informació del Consorci d'Educació de Barcelona	575
Figura 7-1: Objectes descrits en la sèrie documental ESC001-001.....	584
Figura 7-2: Mapa de relacions del procés i la sèrie documental ESC001-001	584
Figura 7-3: Objectes descrits en la sèrie documental GRH023-001	586
Figura 7-4: Mapa de relacions del procés i la sèrie documental GRH023-001	586
Figura 7-5: Objectes descrits en la sèrie documental INF001-001	589

Figura 7-6: Mapa de relacions del procés i la sèrie documental INF001-001	589
Figura 7-7: Objectes descrits en la sèrie documental INF009-001	593
Figura 7-8: Mapa de relacions del procés i la sèrie documental INF009-001	593
Figura 7-9: Objectes descrits en la sèrie documental LOP003-001.....	596
Figura 7-10: Objectes descrits en la sèrie documental LOP003-001.....	596
Figura 7-11: Objectes descrits en la sèrie documental LOP005-001.....	599
Figura 7-12: Mapa de relacions del procés i la sèrie documental LOP005-001	599
Figura 7-13: Objectes descrits en la sèrie documental NEE002-001.....	601
Figura 7-14: Mapa de relacions del procés i la sèrie documental NEE002-001.....	602
Figura 7-15: Objectes descrits en la sèrie documental NEE021-001.....	605
Figura 7-16: Mapa de relacions del procés i la sèrie documental NEE021-001.....	605
Figura 7-17: Objectes descrits en la sèrie documental ORG003-001.....	607
Figura 7-18: Mapa de relacions del procés i la sèrie documental ORG003-001	608
Figura 7-19: Objectes descrits en la sèrie documental ORG004-001.....	610
Figura 7-20: Mapa de relacions del procés i la sèrie documental ORG004-001.....	611
Figura 7-21: Mapa de relacions de sèries documentals amb sistemes d'informació del model relacional.....	622
Figura 7-22: Mapa de relacions de sèries documentals amb sistemes d'informació del model distribuït	623
Figura 7-23: Mapa de relacions de sèries documentals amb sistemes d'informació del model orientat a objectes	623
Figura 7-24: Nivell d'influència de l'índex de transaccionalitat basat en el tipus de dades sobre el model tecnològic.....	635
Figura 7-25: Nivell d'influència del nombre de processos sobre el model tecnològic.	635
Figura 7-26: Nivell d'influència de l'índex de transaccionalitat basat en dades sobre el model tecnològic.....	636
Figura 7-27: Nivell d'influència de l'índex de transaccionalitat basat en objectes sobre el model tecnològic	636
Figura 7-28: Mapa de relacions limitades a l'àmbit funcional directe de la sèrie documental NEE002-001	656
Figura 7-29: Mapa de relacions desenvolupades dels objectes de l'àmbit funcional de la sèrie documental NEE002-001.....	657
Figura 8-1: Aparicions del terme dada a la Llei 39/2015	661
Figura 8-2: Aparicions del terme dada al Decret 76/2020	661
Figura 8-3: Representació de la sèrie documental dins el mapa de coneixement del Consorci d'Educació de Barcelona	688

Índex de gràfics

Gràfic 6-1: Nombre de processos i sèries documentals per família de processos.....	505
Gràfic 6-2: Nombre de sèries documentals pel nombre de relacions amb altres sèries	511
Gràfic 6-3: Percentatge de dades elementals per tipus de dada.....	512
Gràfic 6-4: Percentatge de dades elementals per gran tipus i sèrie documental.....	513
Gràfic 6-5: Percentatge d'ocurrències de dades elementals per tipus d'ocurrència ...	514
Gràfic 6-6: Rangs de valors dels tres índex de dadificació.....	520
Gràfic 6-7: Rangs de valors dels tres índex de transaccionalitat.....	524
Gràfic 6-8: Entorns tecnològics per tipus de procés.....	530
Gràfic 6-9: Tipus de relació entre processos.....	531
Gràfic 6-10: Percentatge de sistemes d'informació per model de servei.....	538
Gràfic 6-11: Nombre de sistemes d'informació per tecnologia.....	540
Gràfic 6-12: Percentatge de sistemes d'informació per model de desenvolupament de la tecnologia.....	542
Gràfic 6-13: Percentatge de sistemes d'informació per model d'emmagatzemament de la tecnologia.....	543
Gràfic 6-14: Evolució del nombre de sistemes d'informació per entorn.....	545
Gràfic 6-15: Nombre de sistemes d'informació posats en producció per any i entorn	546
Gràfic 6-16: Nombre de sistemes d'informació desactivats per any i entorn.....	547
Gràfic 6-17: Nombre de sistemes d'informació per entorn tecnològic i tipus de procés	547
Gràfic 6-18: Percentatge del nombre de vincles de sistemes d'informació amb sèries documentals.....	549
Gràfic 6-19: Perfil tecnològic dels sistemes d'informació vinculats a sèries documentals	549
Gràfic 6-20: Nombre de sistemes d'informació connectats a altres sistemes per entorn tecnològic.....	551
Gràfic 6-21: Proporció d'objectes tipus "cosa" i "acció" per entorn tecnològic.....	556
Gràfic 6-22: Proporció d'objectes tipus "cosa" i "acció" per tecnologia d'emmagatzemament.....	556
Gràfic 6-23: Sistemes d'informació amb major i menor índex d'influència.....	558
Gràfic 6-24: Sistemes d'informació amb el nivell de seguretat més alt del Consorci d'Educació de Barcelona.....	562
Gràfic 6-25: Sistemes d'informació amb el nivell de seguretat baix del Consorci d'Educació de Barcelona.....	562
Gràfic 6-26: Percentatge de dades elementals descriptives de "coses" i "accions" en processos, sèries, sistemes i evidències.....	569
Gràfic 7-1: Formats i mitjans d'autenticació de les evidències-objecte del grup focal	613

Gràfic 7-2: Finalitat en la creació de les evidències-objecte del grup focal	614
Gràfic 7-3: Comparativa de les mètriques d'objectes del grup focal i de la mostra	617
Gràfic 7-4: Comparativa de les mètriques d'objectes del grup focal i de la mostra diferenciades per clúster	618
Gràfic 7-5: Comparativa del nombre de sistemes d'informació per tecnologia i clúster	620

Índex de taules

Taula 2-1: Comparativa dels tres tipus de dades	89
Taula 2-2: Evolució en la responsabilitat sobre les dades	93
Taula 2-3: Models d'arquitectura per a la implantació d'un sistema de gestió de dades mestres.....	101
Taula 2-4: Característiques dels tipus d'expedient.....	126
Taula 2-5: Entorns d'aplicació dels tipus d'expedient	127
Taula 3-1: Pautes d'actuació per al manteniment del mapa d'agents.....	172
Taula 3-2: Pautes d'actuació per al manteniment del mapa de processos	189
Taula 3-3: Pautes d'actuació per al manteniment del mapa de sistemes d'informació	199
Taula 3-4: Criteris per a l'establiment de relacions entre sèries documentals	211
Taula 3-5: Pautes d'actuació per al manteniment del mapa de sèries documentals..	213
Taula 3-6: Pautes d'actuació per a generar un circuit d'expedient únic	246
Taula 3-7: Resultat documental de la transformació a un circuit d'expedient únic.....	247
Taula 3-8: Pautes per a la transformació del circuit documental a datacèntric	263
Taula 4-1: Models de referència per a l'elaboració del model de dades	293
Taula 4-2: Correspondència terminològica entre les entitats del model de dades i els models de referència	431
Taula 5-1: Pes en funció del rol de la sèrie documental	442
Taula 5-2: Càlcul de l'índex d'influència sense i amb factor de correcció	445
Taula 5-3: Pes en funció del tipus i lloc de la ocurrència	447
Taula 5-4: Evidències de la sèrie documental INF001-003	449
Taula 5-5: matriu de variables de la sèrie documental INF001-003.....	449
Taula 5-6: Ròssec individualitzat per dada elemental de la sèrie documental INF001-003	449
Taula 5-7: Tipus de relacions entre tres sèries documentals del procés INF001	451
Taula 5-8: Índex de ròssec intersèrie de les relacions entre tres sèries documentals del procés INF001	451
Taula 5-9: Relacions entre tres sèries documentals i els sistemes d'informació del procés INF001	452
Taula 5-10: Índex de ròssec intersèrie de les relacions entre totes les sèries documentals del procés INF001	452
Taula 5-11: Pes en funció del rol del procés	460
Taula 5-12: Pes en funció del tipus i lloc de la ocurrència	462
Taula 5-13: Pes en funció del tipus de connexió entre sistemes d'informació	471
Taula 5-14: Índex de connectivitat potencial del sistema d'informació SI04.053	472
Taula 5-15: Índex de transaccionalitat a partir del tipus de dades d'alguns sistemes d'informació del Consorci.....	474
Taula 5-16: Pes en funció de l'entorn tecnològic del sistema d'informació	476

Taula 5-17: Pes en funció de l'abast del sistema d'informació	476
Taula 5-18: Índex de rellevància de les dades elementals de la sèrie documental INF001-003.....	482
Taula 5-19: Índex de rellevància de les dades elementals del sistema d'informació SI08.023	484
Taula 5-20: Taula resum de tots els indicadors del sistema de governança de la informació.....	490
Taula 6-1: Canvis que han afectat al Consorci d'Educació de Barcelona durant el projecte.....	497
Taula 6-2: Canvis d'abast i plantejament en la recollida de dades del projecte	497
Taula 6-3: Total d'instàncies per entitat a maig de 2023	498
Taula 6-4: Percentatge de dades recollides per entitat	498
Taula 6-5: Percentatge de processos dadificats per família de processos	499
Taula 6-6: Processos inclosos en el grup focal	500
Taula 6-7: Sèries documentals incloses en el grup focal.....	501
Taula 6-8: Nombre d'ocurrències de dades en sistemes i evidències	502
Taula 6-9: Sèries documentals per productor i tipologia de sèrie	504
Taula 6-10: sèries documentals per tipus de procés	504
Taula 6-11: Nombre de sèries per procés i família de processos	506
Taula 6-12: Mètriques de referència de les sèries documentals del grup focal.....	507
Taula 6-13: Nombre de regles de conservació i eliminació per sèrie documental	509
Taula 6-14: Nombre de regles de conservació i eliminació per tipus de resolució.....	510
Taula 6-15: Nombre de regles d'accés i ús per sèrie documental	510
Taula 6-16: Nombre de regles d'accés i ús per tipus de resolució.....	510
Taula 6-17: Nombre d'objectes per tipus d'objecte i sèrie documental del grup focal	514
Taula 6-18: Sèries documentals amb un major índex d'influència.....	517
Taula 6-19: Sèries documentals amb un menor índex d'influència.....	518
Taula 6-20: Índex de dadificació simple a nivell d'evidència per família de processos	518
Taula 6-21: Índex de dadificació simple a nivell de dada elemental per família de processos	519
Taula 6-22: Índex de dadificació ponderat per família de processos	519
Taula 6-23: Comparativa dels valors dels índex de dadificació per família de processos	520
Taula 6-24: Índex de ròssec intrasèrie per família de processos.....	521
Taula 6-25: Relacions entre sèries documentals amb un índex superior de ròssec intersèrie.....	522
Taula 6-26: Índex de transaccionalitat a partir de les dades elementals per família de processos	523
Taula 6-27: Índex de transaccionalitat a partir del tipus de dades per família de processos	523

Taula 6-28: Índex de transaccionalitat a partir dels objectes per família de processos	524
Taula 6-29: Nombre de processos per tipus de procés	526
Taula 6-30: Nombre de processos per família de processos.....	526
Taula 6-31: Famílies amb un percentatge més alt de processos clau	527
Taula 6-32: Percentatge de tasques automatitzades sobre el total de tasques per família de processos.....	528
Taula 6-33: Percentatges de complexitat i dependència per família de processos....	528
Taula 6-34: Nombre i mitjana d'evidències per família de processos	529
Taula 6-35: Nombre i mitjana de sèries documentals per procés i família de processos	529
Taula 6-36: Nombre d'entitats de dades per procés del grup focal.....	531
Taula 6-37: Índex d'influència per processos del grup focal	533
Taula 6-38: Nombre de sistemes d'informació per entorn	536
Taula 6-39: Nombre de sistemes d'informació per abast.....	536
Taula 6-40: Nombre de sistemes d'informació per model de servei	537
Taula 6-41: Lectura agregada dels models de servei.....	537
Taula 6-42: Nombre de sistemes d'informació per tecnologia	538
Taula 6-43: Tecnologies per model de desenvolupament	540
Taula 6-44: Nombre de sistemes d'informació per model de desenvolupament de la tecnologia	541
Taula 6-45: Tecnologies per model d'emmagatzemament.....	542
Taula 6-46: Nombre de sistemes d'informació per model d'emmagatzemament de la tecnologia	543
Taula 6-47: Nombre de sistemes d'informació per nivell evidencial del contingut.....	544
Taula 6-48: Sistemes d'informació amb més de quatre processos vinculats	548
Taula 6-49: Entorns tecnològics de les sèries documentals del grup focal.....	549
Taula 6-50: Mitjana d'evidències per entorn tecnològic.....	550
Taula 6-51: Sistemes d'informació amb major nombre d'evidències relacionades	551
Taula 6-52: Nombre d'interfícies en funció del tipus i procediment.....	552
Taula 6-53: Nombre d'interfícies en funció del període d'actualització	552
Taula 6-54: Nombre d'interfícies en funció del termini de generació de la versió d'arxiu	552
Taula 6-55: Dades elementals per entorn dels sistemes d'informació	553
Taula 6-56: Dades elementals per abast dels sistemes d'informació	553
Taula 6-57: Dades elementals per model d'emmagatzemament dels sistemes d'informació	554
Taula 6-58: Dades elementals per model de desenvolupament dels sistemes d'informació	554
Taula 6-59: Nombre de dades elementals per tipus de dada i entorn	555
Taula 6-60: Nombre d'ocurrències per caràcter de l'ocurrència i entorn.....	555
Taula 6-61: Relació de sistemes d'informació amb un índex de criticitat més alt	557

Taula 6-62: Relació dels deu sistemes d'informació amb un índex de connectivitat més alt.....	558
Taula 6-63: Índex de transaccionalitat dels sistemes d'informació per entorn	560
Taula 6-64: Índex de transaccionalitat dels sistemes d'informació per abast.....	560
Taula 6-65: Índex de transaccionalitat dels sistemes d'informació per model d'emmagatzemament	560
Taula 6-66: Índex de transaccionalitat dels sistemes d'informació per model de desenvolupament	560
Taula 6-67: Índex de transaccionalitat basat en objectes per entorn.....	561
Taula 6-68: Índex de transaccionalitat basat en objectes per model d'emmagatzemament	561
Taula 6-69: Classificació de seguretat dels sistemes d'informació per abast	561
Taula 6-70: Nombre de dades elementals per tipus de dada	565
Taula 6-71: Dades elementals amb major presència en processos, sèries documentals, sistemes i evidències	566
Taula 6-72: Objectes amb major presència en processos, sèries documentals, sistemes i evidències	567
Taula 6-73: Superobjectes amb major presència en processos, sèries documentals, sistemes i evidències	568
Taula 6-74: Índex de rellevància global de les dades elementals del Consorci d'Educació de Barcelona	570
Taula 6-75: Índex de rellevància global dels objectes del Consorci d'Educació de Barcelona	570
Taula 6-76: Índex de rellevància global dels superobjectes del Consorci d'Educació de Barcelona	571
Taula 7-1: Mètriques dels índex de dadificació i de ròssec de les sèries documentals del grup focal.....	577
Taula 7-2: Anàlisi clúster de les sèries documentals del grup focal.....	580
Taula 7-3: Sèries documentals de la mostra extreta del grup focal	580
Taula 7-4: Circuit evidencial de la sèrie documental ESC001-001	582
Taula 7-5: Sistemes d'informació de la sèrie documental ESC001-001	582
Taula 7-6: Matriu de variables de la sèrie documental ESC001-001	583
Taula 7-7: Circuit evidencial de la sèrie documental GRH023-001.....	585
Taula 7-8: Sistemes d'informació de la sèrie documental GRH023-001	585
Taula 7-9: Matriu de variables de la sèrie documental GRH023-001	585
Taula 7-10: Circuit evidencial de la sèrie documental INF001-001.....	587
Taula 7-11: Sistemes d'informació de la sèrie documental INF001-001	587
Taula 7-12: Matriu de variables de la sèrie documental INF001-001.....	588
Taula 7-13: Circuit evidencial de la sèrie documental INF009-001.....	590
Taula 7-14: Sistemes d'informació de la sèrie documental INF009-001	590
Taula 7-15: Matriu de variables de la sèrie documental INF009-001.....	590
Taula 7-16: Circuit evidencial de la sèrie documental LOP003-001	594

Taula 7-17: Sistemes d'informació de la sèrie documental LOP003-001.....	594
Taula 7-18: Matriu de variables de la sèrie documental LOP003-001	594
Taula 7-19: Circuit evidencial de la sèrie documental LOP005-001	597
Taula 7-20: Sistemes d'informació de la sèrie documental LOP005-001.....	597
Taula 7-21: Matriu de variables de la sèrie documental LOP005-001	598
Taula 7-22: Circuit evidencial de la sèrie documental NEE002-001	600
Taula 7-23: Sistemes d'informació de la sèrie documental NEE002-001	600
Taula 7-24: Matriu de variables de la sèrie documental NEE002-001	600
Taula 7-25: Circuit evidencial de la sèrie documental NEE021-001	602
Taula 7-26: Sistemes d'informació de la sèrie documental NEE021-001	603
Taula 7-27: Matriu de variables de la sèrie documental NEE021-001	603
Taula 7-28: Circuit evidencial de la sèrie documental ORG003-001	606
Taula 7-29: Sistemes d'informació de la sèrie documental ORG003-001.....	606
Taula 7-30: Matriu de variables de la sèrie documental ORG003-001	607
Taula 7-31: Circuit evidencial de la sèrie documental ORG004-001	609
Taula 7-32: Sistemes d'informació de la sèrie documental ORG004-001.....	609
Taula 7-33: Matriu de variables de la sèrie documental ORG004-001	609
Taula 7-34: Nivell de seguretat dels sistemes d'informació del grup focal.....	614
Taula 7-35: Matriu de variables de la sèrie documental GRH023-001	615
Taula 7-36: Matriu de variables de la sèrie documental LOP003-001	616
Taula 7-37: Índex de coincidència sèrie documental - sistema d'informació per model tecnològic	624
Taula 7-38: Sistemes d'informació amb més processos vinculats.....	625
Taula 7-39: Coeficient de correlació del nombre de processos i els models tecnològics	626
Taula 7-40: Processos per sistema d'informació per model d'emmagatzemament ...	626
Taula 7-41: Processos per sistema d'informació per model de desenvolupament	627
Taula 7-42: Caràcter estimat dels sistemes d'informació vinculats a més processos	627
Taula 7-43: Nombre d'evidències per sistema d'informació dels sistemes vinculats a més sèries	628
Taula 7-44: Hipòtesi sobre les relacions entre sistemes d'informació.....	629
Taula 7-45: Models tecnològics dels sistemes d'informació del primer clúster	629
Taula 7-46: Models tecnològics dels sistemes d'informació del segon clúster.....	630
Taula 7-47: Models tecnològics dels sistemes d'informació del tercer clúster	630
Taula 7-48: Sistemes d'informació del tercer clúster que gestionen dades mestres..	630
Taula 7-49: Sistemes d'informació amb un major nombre de coincidències de dades amb altres sistemes	631
Taula 7-50: Sistemes d'informació amb un índex de coincidència major.....	632
Taula 7-51: Sistemes d'informació amb un índex de transaccionalitat basat en dades més alt.....	633
Taula 7-52: Sistemes d'informació amb un índex de transaccionalitat basat en el tipus de dades més alt	633

Taula 7-53: Sistemes d'informació amb un índex de transaccionalitat basat en objectes més alt.....	634
Taula 7-54: Comparativa dels índex de transaccionalitat dels sistemes de caràcter registral.....	634
Taula 7-55: Sèries documentals amb un índex d'influència més alt i més baix	639
Taula 7-56: Sèries documentals del grup focal amb un nombre menor d'evidències	640
Taula 7-57: Sèries documentals del grup focal amb un major índex de dadificació basat en evidències.....	641
Taula 7-58: Sèries documentals del grup focal candidates a sèries recapitulatives ..	642
Taula 7-59: Mitjana d'objectes per tipus d'objecte i tipus de sèrie documental	642
Taula 7-60: Valors dels índex de transaccionalitat per a les sèries documentals del grup focal.....	643
Taula 7-61: Correspondència entre els clusters de transaccionalitat i els tipus de sèrie documental.....	644
Taula 7-62: Correlació entre els índex de transaccionalitats de les sèries documentals i dels sistemes d'informació	645
Taula 7-63: Nombre de visualitzacions de dades per tipus de visualització	646
Taula 7-64: Correspondència entre les visualitzacions de dades i les sèries documentals del Consorci d'Educació de Barcelona.....	646
Taula 7-65: Comparativa entre el diccionari de dades de la sèrie documental NEE002-001 abans i després de la inclusió de la visualització de dades en el circuit evidencial	650
Taula 7-66: Impacte de la inclusió de la visualització sobre l'índex de ròssec intersèrie	652
Taula 7-67: Diccionari de dades de la sèrie documental NEE002-001	653
Taula 7-69: Distribució per objecte i tipus de les dades de la sèrie documental NEE002-001	654
Taula 7-68: Sentències de cada evidència de la sèrie documental NEE002-001	655
Taula 8-1: Proposta d'indicadors per a decidir el model tecnològic per a la dadificació	668
Taula 8-2: Possibles correspondències entre entitats documentals i tecnològiques..	677

Capítol 1. Introducció

1.1 Context

1.1.1 La relació de la gestió documental amb la informació estructurada fins l'explosió del Big Data

La dialèctica sobre si l'objecte de la gestió documental ha d'incloure també la informació estructurada, o limitar-se només als documents en tant que objectes discrets, es remunta a l'aparició de la tecnologia informàtica (Yeo, 2007). A finals de la dècada dels 80, quan es configurava la gestió documental moderna, John McDonald ja plantejava la necessitat d'incloure de forma plena dins l'abast de la gestió documental els sistemes tecnològics que gestionen informació estructurada a les organitzacions (McDonald, 1989).²

En el segon *DLM Forum on electronic records*, celebrat a Brussel·les l'any 1999, Stephen Harries plantejava la dicotomia entre una orientació de la gestió basada en la gestió de documents, que anomenava "*documents model*" (la perspectiva docucèntrica), i una orientació basada preferentment en la gestió d'informació estructurada, és a dir de dades, que anomenava "*database model*" (la perspectiva datacèntrica) (Harries, 2000). Era el moment de l'emergència de la gestió documental electrònica, que intentava dilucidar quina de les dues formes de representar la informació (documents solts o dades en sistemes d'informació) podia garantir millor els principis propis de la diplomàtica que donen valor d'evidència a la informació: la integritat i l'autenticitat. El projecte motor en aquest sentit fou el projecte InterPARES de la Universitat de *British Columbia*.³

L'expressió document-objecte, o *Document-Like Object* (DLO), està reconeguda en l'àmbit de la gestió de metadades per fer referència a qualsevol objecte d'informació que pot ser objecte de metadescripció. Aquest concepte sorgeix en el marc del desenvolupament del model de metadades *Dublin Core*, en el primer *Dublin Core Metadata Workshop* celebrat a Ohio el 1995. Originalment es va definir com a qualsevol entitat o agrupació d'informació digital equivalent a un document tradicional, essencialment textual i amb nombroses propietats en comú amb els documents en paper. El 1996, en el tercer *Dublin Core Metadata Workshop* aquesta definició es va

² Article republicat l'any 2010.

³ El projecte InterPARES va donar origen a una línia de recerca amb un gran impacte dins la disciplina arxivística (Duranti, 1999; Fraser, 1999; Duranti et al., 2002; Magić, 2008; Soler, 2011).

ampliar per incloure qualsevol recurs discret i constant d'informació, que té idèntic contingut per a tots els seus usuaris, que pot formar part d'una col·lecció, i ser metadescribit i recuperat. Actualment el concepte DLO s'utilitza per referir-se a la unitat mínima d'informació digital que pot ser objecte de metadescripció, i inclou text, imatge, àudio i vídeo (Méndez, 2002).

En el context d'aquesta tesi, entenem que la informació es manifesta en forma d'agregats virtuals de dades (dades estructurades dins de sistemes d'informació), o bé en forma d'agregats definitius de dades, dotats d'un contingut, una estructura i un context que en fan una unitat de significació autònoma, reconeixible com a tal de forma natural i formalment independent, la qual anomenarem document (agafant el concepte de document com ens autònom que apareix definit a la norma MoReq2). Entendrem informació i document administratius, els que són objecte de la gestió documental, com aquells que han estat generats o rebuts per una persona física o jurídica, pública o privada, al llarg de la seva existència i en l'exercici de les activitats i les funcions que li són pròpies (agafant com a referència la definició de la Llei 10/2001 d'arxius i documents de Catalunya, en el seu article 2). Del caràcter administratiu de la informació i els documents se n'infereix que tenen valor probatori o evidencial de la forma com es va esdevenir la tramitació que els va generar, i estan per tant investits d'autenticitat diplomàtica.

Tornant a l'evolució de la gestió documental, tres factors es van combinar per donar preponderància a la perspectiva docucèntrica. En primer lloc, l'aparició i popularització dels gestors documentals (*Electronic Document Management Systems*, o EDMS), que es recolzaven en sistemes gestors de bases de dades orientades a objectes. Aquests sistemes es van especialitzar progressivament (*Digital Assets Management, eMail Management, etc.*), i les funcionalitats més específiques de la gestió documental, com la retenció o l'accés, es van anar segregant en mòduls ERMS (*Electronic Records Management Systems*) (Burghof, 2000; McLeod, 2008).

El segon factor fou que la majoria de professionals de la gestió documental, especialment en l'entorn europeu, provinguessin de l'àmbit arxivístic. La similitud del document electrònic, i dels seus mitjans d'autenticació (signatura electrònica basada en certificat), amb el document tradicional en suport paper, combinat en aquell moment amb una manca de formació tecnològica de la professió arxivística, desconixedora dels mecanismes de seguretat propis de les bases de dades, va decantar la confiança dels professionals cap al document-objecte. El document-objecte es va considerar el destí preferent de la informació digital, especialment per a finalitats d'evidència, i també de preservació a llarg termini, i els sistemes d'informació mers generadors de documents-objecte. De fet, es recomanava que els sistemes d'informació generessin sempre una evidència en format de document-objecte quan fos necessari fixar l'autenticitat i la integritat d'una informació digital (Joyce, 2002).

I el tercer factor de preponderància docucèntrica, el qual va consagrar la divergència entre com ha evolucionat la gestió documental a l'àmbit públic i al privat, fou la forma com es va anar configurant normativament, almenys a Europa, l'administració electrònica. Els legisladors europeus, en el moment de regular la validesa dels documents electrònics en uns entorns de gestió que encara eren fortament híbrids (paper-electrònic), van inspirar-se preferentment en el document en suport paper per donar forma als requisits de validesa i gestió del document electrònic (Kampffmeyer, 2000).⁴

Aquesta presa de posició va tenir com a conseqüència l'ajornament en el temps de la reflexió sobre la significació i el valor d'ús de les agrupacions documentals en els entorns de gestió basats en dades. L'assimilació plena, i jurídicament consolidada, del document electrònic vàlid amb el document-objecte, va crear una zona de confort en la qual els professionals de la gestió documental van poder seguir aplicant les seves metodologies de classificació, descripció, accés i conservació sense canvis substancials en relació amb el seu ús en el món paper. Així doncs, quan en la dècada dels 2000 es va iniciar la gran normalització de la gestió documental, amb la publicació de la norma ISO 15489 i totes les seves derivacions (Bustelo, 2012; Cottin & Dessolin, 2012), es va assumir, gairebé com a dogma, que les mateixes metodologies, i els conceptes en els quals se sustentaven, eren aplicables tant a l'entorn paper com al digital, una visió que ha estat recurrent en la bibliografia arxivística.⁵

La gestió documental, en adoptar com a sistemes de referència els EDMS/ERMS, estava prenent com a base d'aplicació unes eines que no s'havien desenvolupat a partir de les necessitats de la professió arxivística, sinó com a resultat de les necessitats de gestió naturals de qualsevol organització, i un temps abans que la professió arxivística

⁴ Especialment il·lustrativa d'aquest pes del concepte de document en suport paper és la regulació de la validesa del document electrònic a l'estat espanyol. La *Ley 39/2015, de 1 de octubre, del Procedimiento Administrativo Común de las Administraciones Públicas*, en el seu article 70.4, exclou explícitament les bases de dades de qualsevol vincle amb les agrupacions documentals ("*Artículo 70.4. No formará parte del expediente administrativo la información que tenga carácter auxiliar o de apoyo, como la contenida en aplicaciones, ficheros y bases de datos informáticas [...]*"). El desplegament reglamentari de les dues Lleis d'Administració electrònica, el *Real Decreto 203/2021, de 30 de marzo, por el que se aprueba el Reglamento de actuación y funcionamiento del sector público por medios electrónicos*, tot i contemplant la possibilitat de conservar bases de dades en arxius electrònics, prioritza novament la visió docucèntrica quan en l'article 54.3 afirma que "*[...] También será posible la inclusión de su información en bases de datos siempre que, en este último caso, consten los criterios para la reconstrucción de los formularios o modelos electrónicos origen de los documentos, así como para la comprobación de la identificación o firma electrónica de dichos datos*". Fins i tot considerant que a la Disposició final segona fa constar la obligatorietat de desenvolupar una *Norma Técnica de Interoperabilidad de tratamiento y preservación de bases de datos*, la qual "*tratará las condiciones y requisitos relativos a la conservación de las bases de datos para garantizar su autenticidad, integridad, confidencialidad, disponibilidad y trazabilidad, y permitiendo la protección, recuperación y conservación física y lógica de los datos y su contexto*".

⁵ La validesa dels mètodes arxivístics aplicats a la gestió dels documents electrònics ha estat estudiada des dels inicis de la digitalització (Bearman, 1991; Bantin, 1998; Menne-Haritz, 2001).

prenguéssim consciència de la seva necessitat (Bak, 2012). Això va donar lloc en els primers temps a moltes implementacions forçades dels instruments arxivístics sobre els EDMS/ERMS, desaprofitant recursos molt més propers a la gestió basada en dades com l'ús avançat de les metadades, l'herència entre objectes, i la possibilitat de combinar el control d'accessos (ACL) amb els registres d'auditoria com a mecanismes per garantir l'autenticitat dins el propi sistema.⁶

Els conceptes de fons, subfons i productor, associats al nivell orgànic, i de sèrie, subsèrie, etc. en el nivell documental es van mantenir doncs vigents sense una lectura crítica de la seva utilitat. Només el criteri funcional va evolucionar, es va sistematitzar (funció, activitat, transacció) i es va assimilar progressivament al concepte de procés, molt més proper a l'entorn empresarial.⁷

Això no fou obstacle per un avenç important de la gestió documental, que es va concentrar en dues línies d'activitat: l'optimització de la implementació dels sistemes en eines EDMS/ERMS (de la mà sobretot de la indústria del *software*)⁸, i la preservació a llarg termini dels documents electrònics. En el primer cas, el *DLM-Forum on electronic records* fou un dels grans impulsors d'aquest canvi a Europa amb la publicació dels *Model Requirements for the Management of Electronic Records (MoReq Specification)* el 2001, *MoReq2* el 2008 i finalment *MoReq2010*. En el segon, és evident que la preservació digital s'ha constituït com una de les disciplines de la gestió de la informació ja consolidades, i present com a procés clau en la majoria d'institucions de memòria col·lectiva (Térmens, 2013).

Aquesta situació de confort no es va mantenir però de forma immutable. Cap a la dècada de 2010, sobretot des de l'àmbit privat, es va començar a qüestionar de forma substancial la utilitat dels sistemes de gestió documental centralitzats i basats únicament en autenticar documents-objecte, i es va començar a incorporar el treball d'altres disciplines a la gestió documental per tal de cobrir un abast més ampli (Bailey, 2009; Bak, 2012; Cumming & Picot, 2014). La crítica al model de gestió documental majoritari incidia especialment en la complexitat de l'adaptació dels instruments arxivístics a la gestió dels documents electrònics,⁹ i en la manca d'una cobertura global, a nivell de

⁶ Encara avui sovintegen implementacions limitades i poc operatives d'eines EDMS/ERMS a causa d'aquesta visió.

⁷ Especialment amb la norma ISO/TR 26122:2008 Information and documentation — Work process analysis for records.

⁸ Un exemple clar d'això són els productes generats arran del tercer *DLM-Forum on electronic records*, sis llibres blancs elaborats per la indústria del *software*, i un model de requeriments funcionals per al desenvolupament i avaluació de solucions de gestió documental: el *Model Requirements for the Management of Electronic Records (MoReq Specification)* (Serra, 2002).

⁹ Greg Bak és una de les veus més contundents dins l'arxivística en qüestionar els quadres de classificació funcionals i les seves implementacions rígides basades en estructures de carpetes (*file plan*), i en apostar per una gestió documental basada en metadades i en classificacions múltiples (Bak, 2012).

control (*compliance*), sobre el conjunt de la informació de l'organització. Alguns estudis (Lappin, 2010) xifraven en un percentatge inferior al 5% el volum d'informació realment controlada pel sistema de gestió documental, i destacaven que el gran gruix que restava fora de control es concentrava especialment en la informació estructurada captiva de sistemes d'informació no connectats amb el sistema de gestió documental, i en els entorns ofimàtics, una part important dels quals també contenia informació estructurada, especialment fulls de càlcul i bases de dades d'usuari final (Smit, Glaudemans, i Jonker, 2017). No era creïble, per tant, que un sistema que només cobria el 5% del volum d'informació digital d'una organització s'atribuís la condició d'exhaustivitat que exigeix la norma ISO 15489 a qualsevol sistema de gestió documental.

Des d'àmbits adjacents a la gestió documental, i especialment des del sector de la indústria i la consultoria, van començar a aparèixer propostes de reforma, algunes certament simplificadores, com l'ús de *Big Buckets*.¹⁰ Una de les propostes amb més impacte fou l'*Information Governance Referent Model* (IGRM)¹¹, juntament amb els principis GARP (*Generally Accepted Recordkeeping Principles*) de ARMA Internacional,¹² les quals feien èmfasi en la necessitat d'integrar la gestió de documents-objecte en el conjunt de la gestió de tota la informació de l'organització. I també es va començar a utilitzar en l'àmbit de la gestió documental un terme d'abast més estratègic que el de gestió, el concepte de **governança** o **govern**.¹³

La professió arxivística va assumir i consolidar aquesta tendència amb la publicació de la família de normes ISO 30300 (Bustelo, 2012), orientades a convertir el sistema de gestió documental en un component estratègic de l'organització, i va començar a treballar per integrar aquesta visió més estratègica i global de governança en la gestió documental (Brooks, 2019).

1.1.2 Aparició del Big Data i els nous models de gestió de la informació estructurada

En paral·lel amb aquesta evolució, la tecnologia directament relacionada amb la gestió d'informació estructurada va fer també un salt molt important. Aquest canvi, considerat

¹⁰ En un intent de simplificar la classificació, la teoria dels *Big Buckets* redueix artificialment el nombre de categories, creant uns contenidors tipus que facilitin l'aplicació de regles de conservació. En aquest cas, la classificació funcional es considera un obstacle que complica innecessàriament als usuaris finals les tasques d'arxiu (Cisco, 2008b).

¹¹ Vegeu <http://www.edrm.net/projects/igrm>.

¹² Vegeu <https://www.arma.org/page/principles>.

¹³ El concepte de Govern (integral) de la informació fou plantejat per primer cop l'any 2011 com a part de l'*Information Governance Reference Model* (IGRM) de EDRM.net. Per a l'origen del concepte vegeu Kooper, Maes, i Lindgreen, 2011.

per alguns autors una revolució (Gandomi & Haider, 2014), es resumeix sovint fent referència al concepte de *Big Data*, però en realitat té un abast molt més ampli, que es desplega bàsicament en tres vessants.

El primer és l'increment en la capacitat de capturar informació, sovint de forma desatesa, cosa la qual facilita que es capturi en grans volums (Gandomi & Haider, 2014; Mayer-Schönberger & Cukier, 2014). L'ús de dispositius mòbils permanentment geolocalitzats, i integrats de forma transparent en l'hàbitat de l'usuari final (Internet de les coses, *wearables*, sensors, *fog computing*, etc.) ha multiplicat la quantitat d'informació que és possible capturar. Aquesta informació es pot emmagatzemar en molts formats, però la seva captura, majoritàriament automatitzada, facilita que sempre estigui metadescripta, i que molt sovint s'emmagatzemi en forma estructurada. L'evolució dels formats (aparentment no estructurats però basats en XML, JSON, etc.) també proporciona cada cop més informació estructurada (Bellatreche et al., 2021).

El segon vessant és l'increment en la capacitat de processar dades de forma massiva, el nucli del que s'ha anomenat *Big Data*. L'aparició de les bases de dades NoSQL (*Not Only SQL*) i les arquitectures *Hadoop* han fet possible realitzar de forma ràpida operacions de tractament sobre conjunts d'informació heterogenis, parcialment estructurats o no estructurats de forma homogènia. A despit de sacrificar precisió, la possibilitat d'obtenir en el tractament de les dades un rendiment superior per al mateix volum al que poden assolir les arquitectures relacionals tradicionals ha fet molt més assequible la generació de productes d'informació que fins fa un temps estaven reservats a un àmbit especialitzat (Chen, 2022; Wieder & Nolte, 2022).

I són precisament aquests productes el tercer vessant on s'ha produït la disrupció. La presentació visual, dinàmica i interactiva de les dades s'ha convertit en la forma habitual de consumir-les, i en un entorn d'integració sobre el qual és possible prendre decisions (Lemieux, Gormly, i Rowledge, 2014; Lemieux, 2015; Joseph & Hartel, 2017). I cada cop més són màquines qui prenen aquestes decisions mitjançant algorismes d'aprenentatge (*machine learning*) i predicció (intel·ligència artificial) (Rolan et al., 2019).

L'impacte sobre la gestió documental no és menor. L'acceptació social de la presa de decisions a partir de visualitzacions, o la presa de decisions automatitzada a partir de dades tractades algorítmicament, comporta un desplaçament de la confiança col·lectiva en l'autenticitat i la integritat de la informació dels documents cap a les dades (Lemieux, Gormly & Rowledge, 2014), aquelles que habitualment s'anomenen "dades en brut". Igualment s'esdevé amb l'acceptació acrítica de les dades obertes, on el potencial de reutilització les fa preferibles a fonts originals, menys tractades però al mateix temps més difícils de processar (Borglund & Engvall, 2014; Parycek, Höchtl, i Ginner, 2014; Shepherd et al., 2019). Aquest impacte s'ha vist amplificat per la crisi dels mecanismes d'autenticació i signatura basats en certificat electrònic¹⁴, i per l'aparició, des de l'àmbit

¹⁴ Crisi bàsicament provocada per la baixa implantació social de l'ús de la signatura electrònica avançada o qualificada basada en certificat, que ha portat a la legislació a flexibilitzar

del *gaming* i de les criptomonedes, de la tecnologia de blocs (*Blockchain*), que ha estat ràpidament acceptada i incorporada pel sector privat (*Smart Contracts*) (Alamillo et al., 2021).

Aquest no ha estat l'únic element d'impacte sobre la gestió documental i l'arxivística. L'arquitectura tradicional amb la qual les organitzacions estructuraven els seus sistemes d'informació (aplicacions verticals, sistemes de gestió de dades mestres, sistemes d'explotació o *business intelligence*), amb la qual es podia integrar de forma relativament ordenada un EDMS/ERMS, també ha evolucionat cap a models cada cop més flexibles (*Data Lake* o *Data Lakehouse*) (Khine & Wang, 2018; Sawadogo & Darmont, 2021; Wieder & Nolte, 2022) i de gestió descentralitzada (*Data Mesh*) (Delgado, 2022), i amb un impacte acusat en la capacitat de l'usuari final de desenvolupar sistemes d'altres prestacions a través de models *SaaS*, com *Google Drive*, *Microsoft Sharepoint*, *Alfresco Cloud*, etc. (Ghiran & Buchmann, 2019; Lappin et al., 2021).

A diferència del que es podia pensar, lluny de potenciar la idea d'un govern integrat de la informació en la línia del IGRM o de ISO 30300,¹⁵ aquesta anomenada "Revolució del Big Data" s'ha anat configurant com una disciplina amb entitat pròpia, el *Data Science*.¹⁶ Incorporant tècniques d'altres disciplines com l'estadística (analítica, predictiva, etc.) o la informàtica, la ciència de les dades evoluciona donant origen a nous conceptes (l'inatge de la dada, catàleg de dades, etc.) i a noves figures professionals (*Data Architect*, *Data Steward*, etc.). I en els darrers anys, amb un impacte en l'imaginari de les organitzacions que porta a considerar el govern de la dada com el pal de paller de la gestió de la informació (Khatri & Brown, 2010; Benfeldt, Persson, & Madsen, 2020; Järvihaavisto & Öhman, 2022; Mordell, 2022).

progressivament els mitjans d'autenticació i signatura admesos com a vàlids en les transaccions amb el sector públic (clau concertada, combinació clau-dispositiu mòbil, etc.) (Alamillo, 2021).

¹⁵ Per als models de referència per a la governança de la informació vegeu l'apartat 3.1.2 del capítol 3.

¹⁶ S'ha evolucionat de referir-se al tractament i explotació de dades com a analítica de dades a anomenar-lo "ciència de dades", especialment pels seus vincles amb la matemàtica i l'estadística. De fet, el concepte "ciència de dades" el va utilitzar per primera vegada el científic danès Peter Naur a la dècada dels seixanta com a terme alternatiu a ciències de la computació, i posteriorment, en el context d'una reunió de la *International Federation of Classification Societies* (IFCS) a Kobe, el concepte fou associat a la classificació i a l'estadística. Però no va ser fins el 2001 que l'estadístic William S. Cleveland va proposar la ciència de dades com a disciplina independent. Paradoxalment, aquest caràcter científic no s'ha atribuït a la gestió documental, malgrat la seva vinculació amb la matemàtica o la computació en els sistemes de classificació i descripció.

1.1.3 Conseqüències sobre la gestió documental i l'arxivística

La gestió documental i l'arxivística, immerses en una evolució cap al govern integrat de la informació (model ISO 30300), han reaccionat tard, i fins el moment de forma no massa efectiva, a la "Revolució del Big Data".

El corrent de les dades obertes, especialment en l'àmbit públic (*Open Government Data*) fou un dels primers àmbits d'on els arxius, responsables naturals de l'accés a la informació, quedaren desplaçats durant un temps (Parycek, Höchtl, & Ginner, 2014; Borglund & Engvall, 2014; Hellberg & Hedström, 2015; Attard et al., 2015; Shepherd et al., 2019). A títol d'exemple, pràcticament cap portal de dades obertes de l'estat espanyol, malgrat oferir conjunts de dades (*datasets*¹⁷) en un context de transparència administrativa i amb presumpció de veracitat, és a dir considerats com autèntics, fou iniciativa del servei d'arxiu ni tingué en origen relació directa amb els arxius, malgrat ser en essència informació purament d'arxiu, emanada de forma original dels sistemes d'informació de gestió.¹⁸ Tampoc es van utilitzar inicialment en les iniciatives *Open Government Data* models de descripció o metadades d'origen arxivístic per a descriure els *datasets* (Khatri & Brown, 2010; Ehrlinger et al., 2021). Certament els arxius han incorporat progressivament els *datasets* dins els seus arxius digitals, però competint amb nombrosos arxius paral·lels (*Data archives*, *Data catalogs*, etc.) d'origens ben diversos (Franco & Pérez, 2015; Thomson, 2016; Shepherd et al., 2019).

La recuperació de terreny per part de l'arxivística i la gestió documental ve de la mà, per una part, de la progressiva incorporació dels *datasets* en els arxius digitals i repositoris de preservació digital (Shepherd & Smith, 2000; Niu, 2016). A partir de la competència dels arxius en matèria d'accés i transparència, també s'han vinculat progressivament als repositoris de dades obertes, amb projectes comuns (Casellas, 2014). Un altre àmbit on els arxius estan avançant, sobretot amb els treballs de l'escola de Vancouver, és en l'acceptació de *Blockchain* com a mitjà d'autenticació, i en explorar el seu potencial com a possible entorn d'arxiu (Lemieux, 2016; Findlay, 2017; Hofman et al., 2019; Duranti, 2022).

Relacionada a aquesta tendència es troba també la idea dels arxius distribuïts, vinculats a través de *Blockchain*, i començant a trencar la idea de fons i productor (Franks, 2020). També són importants els avenços en els plantejaments ontològics dels models de descripció arxivístics, tenint com a principal exponent el model *Records in Context* del

¹⁷ En l'estat de la qüestió es detalla el debat sobre la definició i significació del concepte de *dataset*.

¹⁸ Amb algunes excepcions, com l'Ajuntament de Girona (Casellas, 2014) o el Govern d'Aragó (Akerreta & Moyano, 2021).

Consell Internacional d'Arxius.¹⁹ I també els projectes que incardinen en els aspectes arxivístics de la intel·ligència artificial i els biaixos en els algoritmes que puguin afectar l'autenticitat de la informació (Frontoni et al., 2022; Duranti, 2022).

El principal buit el trobem tanmateix en el camp de la gestió documental. Hi ha poca bibliografia que reflexioni sobre els mètodes normalitzats de la gestió documental, tant a nivell de processos (captura, descripció, classificació, tria i preservació) com d'instruments (quadres de classificació, regles de conservació i eliminació, etc.) en l'intent de vincular-los a elements cabdals del govern de la dada com són el control del llinatge de la dada, la significació documental dels *datasets* i dels sistemes d'informació que els emmagatzemen i produeixen, i el valor arxivístic de les visualitzacions de dades.

Aquesta absència allunya la gestió documental del govern de la dada, i minora progressivament la seva participació en el govern integrat de la informació de l'organització (Lappin, 2010). La ciència de dades va camí d'incorporar conceptes, com el de l'autenticitat vinculada al llinatge, el control de l'accés, o la tria, conservació i eliminació de les dades, que són propis de la gestió documental, però també de fer-ho sense la participació activa de la gestió documental més enllà d'aportacions puntuals.²⁰ I bàsicament perquè la gestió documental no està fent evolucionar els seus conceptes, i especialment el d'agrupació documental (clau per a l'aplicació de polítiques d'accés i conservació) en la línia de fer-los aplicables a les formes pròpies de la gestió basada en dades.

Aquest és precisament l'objectiu que es planteja aquesta tesi. Una aplicació informàtica, i la base de dades que li dona suport, és un constructe tecnològic. La pregunta que es planteja la tesi és si els límits de l'aplicació coincideixen amb els d'una agrupació documental, i si aquesta associació permet que l'aplicació (i les dades que genera i gestiona) heretin de forma directa les polítiques assignades a aquella agrupació documental. El que es vol escatir és, doncs, quina part o manifestació d'un entorn de treball dadificat pot adquirir significació documental i ser objecte de l'aplicació de polítiques de gestió documental, i també si aquesta significació pot venir donada per la pertinença lògica a una agrupació documental tradicional.

¹⁹ Els plantejament ontològics aplicats a la gestió documental no són nous (Klischewski, 2003), però estan assolint un protagonisme creixent sobretot arran de la nova norma de descripció arxivística *Records in Context* del Consell Internacional d'Arxius.

²⁰ *DAMA International*, l'associació internacional de gestió de dades (<https://www.dama.org/cpages/home>) ha condensat l'essencial de la gestió de dades en onze principis o àmbits d'actuació, els quals ha recollit en la segona edició de la monografia *Data Management Body of Knowledge (DMBOK2)*. D'aquests onze principis, només un contempla explícitament la gestió documental, malgrat que en bona part de la resta s'hi poden trobar també vincles amb eines i mètodes propis de la gestió documental (Serra, 2021). I quelcom similar s'esdevé amb la norma *ISO/IEC 38505-1:2017 Information technology — Governance of IT — Governance of data — Part 1: Application of ISO/IEC 38500 to the governance of data*.

1.2 Justificació de l'objecte i interès per a la disciplina arxivística

Les agrupacions documentals, segons les normes de descripció arxivística vigents, són la base per l'aplicació de polítiques de gestió documental.²¹ Aquestes polítiques no s'apliquen mai al nivell d'unitat documental, sinó d'agrupació documental (MoReq2010). La pervivència de la gestió documental com a disciplina instrumental depèn doncs fortament del concepte d'agrupació documental.

Les organitzacions gestionen la seva informació en forma d'unitats d'informació. Quan aquestes unitats d'informació tenen la forma d'objectes, es poden agrupar en agrupacions documentals, i es poden automatitzar d'acord amb un model molt similar a l'aplicat anteriorment sobre paper. Però quan aquestes unitats d'informació tenen la forma d'agregats circumstancials de dades, no s'emmagatzemen com objectes, sinó com dades distribuïdes dins de sistemes d'informació o bases de dades. I les bases de dades responen a una estructura purament tecnològica, sense *a priori* cap significació documental.

La pregunta que es planteja aquesta tesi és si, en un escenari fortament dadificat, on la majoria d'unitats d'informació estan compostes de dades repartides entre diferents bases de dades o sistemes d'informació, es pot seguir aplicant el concepte d'agrupació documental. I si la resposta és afirmativa, de quina forma pot seguir-hi existint i operant aquest concepte, i amb quina aportació de valor.

Una resposta afirmativa permetria replantejar l'actual orientació a considerar els conjunts de dades (*datasets*) entitats documentalment significatives, tal com apunten les iniciatives de preservació de conjunts de dades,²² i permetria desplaçar el focus d'atenció cap al control de la creació d'aquests conjunts de dades, el qual és l'àmbit natural d'actuació de la gestió documental. Això reforçaria la participació de la gestió documental en la gestió de dades, i facilitaria la integració d'ambdues visions en un context de govern integral de la informació, una tesi defensada pels autors que aposten més decididament per la idea d'integrar la gestió documental en un únic sistema de govern integral de la informació (Lueders, 2021; Kampffmeyer, 2023; ISO/TC46/SC11-Archives/records management, 2023).

Atesa la migradesa de referents pràctics en l'orientació a dades dins la disciplina arxivística, s'ha optat en aquesta tesi per utilitzar una metodologia de caràcter empíric, amb la voluntat d'elaborar el mapa de la informació d'una organització estàndard (en el

²¹ Per al detall de les normes, vegeu l'apartat 4.5 del capítol 4.

²² Els *Data Management Plan* propis dels projectes de recerca quantitativa són un exemple d'aquesta orientació a la preservació de *datasets* com a producte final (Redkina, 2019).

nostre cas del sector públic) al nivell de dada, posant a prova un mètode de treball que pugui ser d'utilitat per a l'experimentació i la recerca en d'altres projectes similars.

1.3 Hipòtesis i objectius

Els objectius d'aquest projecte són de dos tipus: **teòrics** i **instrumentals**. Procedim a detallar-los.

1.3.1 Objectius teòrics

L'objectiu principal és analitzar en quina mesura el concepte d'agrupació documental amb significació arxivística, i en especial el concepte de sèrie documental (i el seu derivat natural l'expedient), és present en la tramitació electrònica, sobretot en aquella tramitació on la major part de la informació que es genera, si no tota, està en format de dades estructurades i no de documents-objecte.

Aquest objectiu general s'ha estructurat en set objectius específics. Els dos primers tenen caràcter descriptiu, i els cinc següents tenen caràcter explicatiu. Els objectius específics són els següents:

1.3.1.1 Objectiu 1: característiques de la sèrie documental dadificada

Aquest objectiu pretén conèixer, a partir de casos d'aplicació, quines són les característiques d'una sèrie documental que ha passat per un procés de dadificació.²³ Aquesta descripció es vol elaborar a partir d'una lectura ampliada que permeti extreure'n conclusions de caràcter general, qüestionant la lectura tradicional del concepte de sèrie documental mitjançant la seva inclusió dins un marc més ampli, que implicarà:

- Les característiques inherents al circuit evidencial de la sèrie documental: formats, mitjans d'autenticació, i nivell de dadificació que es pot assolir.
- Les característiques del diccionari de dades i de la matriu de variables de la sèrie documental relatives a les dades elementals i als objectes descrits.
- Les relacions entre les sèries documentals en el marc dels processos que les produeixen.
- La identificació i anàlisi de les relacions entre les sèries i els sistemes d'informació, aspecte que es desenvolupa descriptivament en l'objectiu 2 i analíticament en l'objectiu 3.

²³ Per conèixer les característiques d'aquest procés vegeu l'apartat 3.3 del capítol 3.

1.3.1.2 Objectiu 2: característiques de l'ecosistema tecnològic de la sèrie documental dadificada

Com a continuació de l'objectiu 1, es vol analitzar quines formes tecnològiques pren la sèrie documental dadificada, en quin tipus de tecnologia es materialitza, i quins elements componen l'expedient electrònic dadificat dins aquestes tecnologies.

Com a part d'aquesta lectura, es vol analitzar de quina forma s'estructuren els sistemes d'informació a les organitzacions per a contenir les sèries documentals dadificades.

Sobre la base d'aquest coneixement, es planteja la verificació o refutació de cinc hipòtesis, les tres primeres explicatives (objectius 3, 4 i 5) i les dues darreres amb caràcter de proposta metodològica (objectius 6 i 7).

1.3.1.3 Objectiu 3: equivalència entre la sèrie documental i el sistema d'informació

La hipòtesi inicial d'aquest objectiu és que els sistemes d'informació d'una organització són equiparables a agrupacions documentals, i específicament al concepte de sèrie documental (és a dir, es compleix l'equiparació sèrie-sistema), i per tant els sistemes d'informació en el seu conjunt poden ser objecte, amb les degudes adaptacions, del tractament propi de la gestió documental (conservació/eliminació i accés). Per tant, s'infereix que el concepte de sistema d'informació (de gestió o de negoci) comparteix límits conceptuals amb les agrupacions documentals arxivístiques, i té per tant significació documental.

La confirmació d'aquesta hipòtesi obre la porta a que la gestió documental apliqui polítiques directament sobre els sistemes d'informació, i consideri les dades de cada sistema com un conjunt amb significació documental (equivalent a sèrie documental).

1.3.1.4 Objectiu 4: les arquitectures distribuïdes com a entorns de gestió (transaccionals)

La hipòtesi d'aquest objectiu és que les arquitectures distribuïdes, representades actualment pels models *data lake* i *data fabric*, són l'entorn preferent de gestió (l'entorn de tramitació o transaccional), mentre que els sistemes verticals i basats en models relacionals són sistemes orientats a la consolidació de dades mestres.

Aquesta hipòtesi és directament oposada al corrent majoritari, que considera les aplicacions convencionals, basades en bases de dades relacionals, com els sistemes

de tramitació per excel·lència (els eixos funcionals d'una organització), i els entorns distribuïts com entorns orientats bàsicament a l'emmagatzematge de dades amb finalitats d'explotació i analítica.

En aquesta equació s'introduiran dos elements que es poden combinar amb els entorns distribuïts: les plataformes de gestió documental, i els desenvolupaments *low code*. I s'afegirà també l'ús de visualitzacions de dades per a la gestió, que s'analitza en l'objectiu 6.

1.3.1.5 Objectiu 5: equivalència entre tipus de sèries documentals i tipus de sistemes d'informació

La hipòtesi d'aquest objectiu és que l'equivalència entre sèrie documental i sistema d'informació ve determinada per la tipologia de cadascuna de les dues entitats. Així doncs, les sèries documentals de tramitació són equivalents a sistemes d'informació transaccionals²⁴, mentre que les sèries documentals recapitulatives s'associen a sistemes de gestió de dades mestres.

Amb aquesta hipòtesi es vol explorar la possible vinculació entre els tipus de sèrie documental (i les relacions que mantenen entre elles), i l'arquitectura de sistemes d'informació de l'organització, amb el supòsit metodològic que, a través del mapa de sèries documentals, es puguin fer aportacions a la planificació i racionalització d'aquesta arquitectura de sistemes.

1.3.1.6 Objectiu 6: les visualitzacions de dades com a recurs per a la gestió

Tradicionalment les visualitzacions dinàmiques de dades s'han vist com quelcom vinculat de forma gairebé exclusiva a l'explotació de la informació amb posterioritat a la gestió, o com a eina de control a nivell de comandament. Però l'emergència d'una generació d'eines de visualització prou usables per posar-les a l'abast dels usuaris finals ha obert la porta a donar-los nous usos en la gestió.

La hipòtesi d'aquest objectiu és que la inclusió sistemàtica de visualitzacions de dades com a recurs per a la gestió modifica les característiques de la sèrie documental i la seva integració i rol en el mapa de sèries de l'organització, i reforça, en cas de verificar-se, el que s'afirma en l'objectiu 4: que les arquitectures distribuïdes tipus *data lake* i *data fabric* poden actuar com a sistemes o entorns de gestió.

²⁴ Els que habitualment s'anomenen sistemes de gestió.

1.3.1.7 Objectiu 7: de la classificació funcional a la classificació ontològica

Finalment, aquest objectiu pretén comprovar si, en un entorn dadificat, és l'òptim seguir aplicant una classificació funcional i jeràrquica, d'acord amb el model ISO 15489, com la que s'aplica actualment de forma majoritària als arxius administratius catalans. En aquest sentit, es valoraran tres alternatives:

1. El criteri funcional i l'estructura jeràrquica són aplicables a un entorn dadificat (i per tant permeten classificar sistemes d'informació o conjunts de dades).
2. El criteri funcional i l'estructura jeràrquica no són aplicables a un entorn dadificat, on es poden substituir per models ontològics amb criteris no necessàriament funcionals.
3. El criteri funcional i l'estructura jeràrquica són compatibles amb models ontològics i criteris no funcionals en un mateix entorn de gestió dadificada.

1.3.2 Objectius instrumentals

Per poder avaluar el compliment de les hipòtesis calia disposar d'una imatge el més completa possible del mapa informacional d'una organització, en un entorn on tant el sistema de gestió documental com una governança integral dels sistemes d'informació i de les dades que aquests contenen estigués completament implantat. Calia disposar en conseqüència de mapes complets de processos, agrupacions documentals, sistemes d'informació i dades. Aquesta necessitat ha implicat desenvolupar i implantar un sistema de govern de la informació en un ens (el que ha estat objecte d'estudi), i desenvolupar una metodologia que permetés agilitzar-ne la implantació en un temps prou curt, amb l'objectiu de poder disposar d'una base d'informació per a l'anàlisi i confirmació o desestimació de les hipòtesis plantejades. L'ens triat ha estat el Consorci d'Educació de Barcelona, les característiques del qual es descriuen en l'apartat 6.1 del capítol 6.

Donat el baix nivell inicial de digitalització de l'ens del cas d'aplicació, ha calgut també transformar sistemes existents per tal d'assolir el nivell més alt possible de dadificació²⁵ i poder disposar d'un mapa de dades prou ampli per a l'anàlisi.

Aquesta necessitat, la de generar un model exhaustiu, no només d'identificació i descripció sinó també de transformació, per a la recollida d'informació, ha obert la porta a plantejar-se la utilitat d'aquest model com a sistema de govern de la informació (a

²⁵ El concepte de dadificació s'explica detalladament en la descripció d'un dels instruments que s'ha elaborat, la Guia de dadificació de processos (**GDAD**) (vegeu l'apartat 3.3 del capítol 3).

nivell merament descriptiu) aplicable a d'altres organitzacions de característiques i dimensions similars. D'aquesta constatació s'han derivat quatre objectius instrumentals:

- Validar la utilitat de la metodologia per a la dadificació de processos com a factor accelerador en la transformació digital de les organitzacions.
- Validar la facilitat d'ús i comprensió de la metodologia per a la identificació i mapatge utilitzada per a la recollida de les dades per part d'equips sense un coneixement previ.
- Validar el nivell d'esforç necessari per a recollir i mantenir actualitzada la informació descriptiva i de mapatge requerida per la metodologia.
- Validar la utilitat dels mapes i descripcions per a la localització de fonts d'informació i dades dins el conjunt de l'organització per part d'equips d'anàlisi i avaluació.

Els objectius es plantegen tant dins l'àmbit dels diferents agents especialitzats d'una organització (responsables de gestió documental i arxiu, gestors de tecnologies de la informació, oficines de dades, etc.) com en l'àmbit de la prestació de serveis externs (consultoria).

A partir de les quatre validacions indicades s'ha intentat derivar un conjunt de recomanacions per a l'aplicació de la metodologia proposada en una organització.

1.4 Límits i abast

L'àmbit d'estudi i aplicació del projecte és molt ampli, i no és voluntat ni està a l'abast de l'autor tractar-lo en tota la seva extensió i complexitat. Per assegurar la factibilitat de l'estudi s'ha dimensionat establint els següents límits en el seu abast:

- A. La recerca bibliogràfica s'ha centrat en les metodologies i models de referència de cada disciplina, i especialment en aquells consolidats a través de normes tècniques i estàndards. S'ha considerat també aquelles fonts bibliogràfiques que feien aportacions més rellevants a nivell conceptual o de model de gestió, sense analitzar amb detall totes les seves variants ni les aportacions més empíriques.
- B. La recerca ha fet especial èmfasi en les fonts i models de referència d'àmbit estatal, i en la tradició arxivística catalana i espanyola, pel seu impacte directe en el model que s'ha treballat en la fase experimental.
- C. En tots els casos s'ha prioritzat la visió emanada de l'arxivística i la gestió documental, i especialment en aquells àmbits objecte d'aquesta disciplina (la creació, l'autenticació, l'ús i la preservació). Les visions pròpies d'altres disciplines (gestió de processos, ciència de dades, etc.) s'han incorporat només en allò que feia referència als processos de gestió documental ja esmentats.
- D. La part experimental s'ha portat a terme en una única organització, el Consorci d'Educació de Barcelona, la qual per les seves dimensions i estat inicial ofería un camp de treball òptim per al desenvolupament del projecte en totes les seves facetes. Només en determinats casos la metodologia s'ha contrastat amb aplicacions puntuals en altres organitzacions.²⁶
- E. L'organització triada també determina que el model sigui aplicable preferentment en l'àmbit públic, o en organitzacions productores de serveis, que poden ser de l'àmbit públic o privat. No es valora per tant l'aplicabilitat del model a entorns industrials o de producció de béns.

²⁶ En el capítol 6 es descriuen les característiques i magnituds del Consorci d'Educació de Barcelona i la seva situació en el moment de començar el projecte.

- F. El model experimental proposat és un model bàsicament conceptual, i no s'ha plantejat tampoc cap implementació tecnològica específica. També s'ha descartat incorporar una implementació ontològica, tot i que possible, bàsicament per tres motius:
- a. Perquè els sistemes de governança de la informació es construeixen a mida de cada organització, i no es planteja en un principi la seva compartició pública.
 - b. Perquè són sistemes extremadament dinàmics, que evolucionen i es modifiquen de forma gairebé constant.
 - c. Per les pròpies limitacions d'esforç i abast del projecte.

1.5 Metodologia

L'estudi té caràcter de recerca avaluativa de caràcter qualitatiu. El desenvolupament del projecte s'ha estructurat en cinc fases, les quals es detallen a continuació.

1.5.1 Fase 1: Anàlisi de l'estat de la qüestió

El primer pas ha estat l'anàlisi de l'estat de la qüestió. L'abast bibliogràfic del projecte era ingent, tant per la generalitat dels conceptes inclosos com pel fet d'ubicar-se en la intersecció de quatre disciplines: l'arxivística i la gestió documental, la gestió de processos, la informàtica, i la ciència de dades. Un segon factor de complexitat ha estat el fet que, degut a la necessitat de treballar sobre un sistema de gestió substancialment dadificat, el qual en bona part s'ha hagut de construir de bell nou, el projecte s'ha dilatat molt en el temps. Així doncs, a la transversalitat entre disciplines s'ha sumat l'evolució d'aquestes al llarg del temps, cosa la qual ha multiplicat les fonts documentals a considerar.

Per aquest motiu, s'ha optat per una cerca selectiva de les referències de cada disciplina i àmbit temporal. S'ha pres com a referència aquell coneixement consolidat en cada disciplina a través de les normes tècniques, i s'ha complementat amb una cerca en bases de dades bibliogràfiques i referències creuades dels articles més rellevants pel que fa a les aportacions conceptuals i a la reflexió sobre cada disciplina.²⁷ En determinats casos, especialment amb els conceptes clau (sèrie documental, *dataset*, etc.) s'ha realitzat una cerca exhaustiva dels articles publicats els darrers cinc anys en les principals publicacions científiques de la disciplina que facin esment d'aquests termes, amb importants buits detectats, els quals es detallen en l'estat de la qüestió.

Els eixos temàtics d'aquesta cerca han estat els següents:

- L'evolució de la gestió documental cap al concepte de govern integrat de la informació.
- L'evolució dels models de descripció propis de l'arxivística i la gestió documental.

²⁷ La cerca s'ha fet utilitzant preferentment els recursos del Centre de Recursos per a l'Aprenentatge i la Investigació de la Universitat de Barcelona (CRAI-UB, disponible a <https://crai.ub.edu/>), a través del qual s'ha accedit a diverses bases de dades bibliogràfiques. També s'han utilitzat repositoris d'accés obert com *e-prints in library & information science* (e-Lis, disponible a <http://eprints.rclis.org/>), Revistes catalanes amb accés obert (RACO, disponible a <https://www.raco.cat/>) i la xarxa *ResearchGate* (<https://www.researchgate.net/>), juntament amb blocs d'experts en gestió documental i recursos del món empresarial i dels productors de solucions.

- Els models de descripció d'entitats i conjunts de dades propis de la ciència de dades.
- L'evolució i vigència actual del concepte de sèrie documental i els seus derivats.
- L'evolució dels conceptes funcionals (procés o activitat) que donen origen a les sèries documentals.
- El concepte de sistema d'informació en tant que entorn tecnològic que genera i emmagatzema dades.
- El concepte de dada, les seves manifestacions agregades i la dialèctica amb el document.
- Metodologies per a la governança de la dada i de la informació en general i estudis sobre recerques similars.

A partir del buidatge d'aquestes fonts s'ha redactat l'estat de la qüestió i s'ha donat forma a la metodologia que s'ha desenvolupat per a la recollida i anàlisi de la informació.

1.5.2 Fase 2: Preparació i recollida d'informació

Per poder analitzar les relacions detallades en els objectius calia disposar del mapa complet d'actius d'informació d'una organització, i especialment de les relacions entre els seus components fins al nivell de dada. L'estat de la qüestió ha posat de relleu que, malgrat les nombroses eines que aporta cada disciplina per al seu àmbit de coneixement, manquen referents metodològics que aportin una visió completament integrada des del punt de vista operatiu, i que incorporin en tots els casos la gestió a nivell de dada (Kampffmeyer, 2023). Per aquest motiu, el primer pas ha estat fixar uns criteris metodològics i instrumentals propis per a la construcció de l'esmentat mapa.

1.5.2.1 Fixació dels criteris metodològics d'identificació

Amb l'objectiu de fixar aquests criteris s'ha elaborat dues guies metodològiques. La primera guia, que hem anomenat **Guia d'inventari i mapatge (GMAP)**, segueix un plantejament *top-down*, i té per objectiu facilitar els criteris i el model descriptiu per identificar les diferents entitats a relacionar (bàsicament processos, sèries, sistemes d'informació i dades), a partir d'un treball de camp d'inventari i descripció, i a continuació construir el mapa de relacions i la base per a l'anàlisi posterior. Aquesta primera guia

s'inspira en els conceptes derivats dels models de descripció arxivístics, i especialment en aquells de base ontològica.²⁸

La segona guia, que hem anomenat **Guia de dadificació de processos (GDAD)**, segueix un plantejament *bottom-up*, i respon a la necessitat, un cop construït el mapa de relacions, de treballar a nivell operatiu, identificant totes les entitats implicades en un mateix procés o activitat, novament fins al nivell de la dada. Aquesta operació, repetida per cadascun dels processos, alimenta en una segona capa el mapa global d'informació de l'organització, i permet completar l'anàlisi amb informació de detall sobre l'ús de cada dada en el procés. Aquesta segona guia s'inspira en la metodologia *Administrative Process Study (APS)*,²⁹ complementada amb aportacions pròpies en la part corresponent a la dadificació d'un procés i la seva integració amb el mapa de dades global.

Juntament amb les guies metodològiques, i amb l'objectiu d'estructurar la informació recollida, s'ha elaborat un model de descripció propi. Aquest model s'ha construït a partir del buidatge dels principals models de descripció arxivístics i de l'àmbit de la gestió documental, com també de models utilitzats per elaborar catàlegs de dades.

1.5.2.2 Desenvolupament del sistema de recollida i anàlisi

Per a la recollida de la informació i la construcció dels mapes s'ha desenvolupat una eina informàtica específica. Aquesta eina es va desenvolupar inicialment com un entorn d'usuari final en modalitat *SaaS*, basat en *Microsoft Sharepoint*, a partir d'una plantilla prèvia de recollida en format *Microsoft Excel*. Aquesta eina es va anomenar **CATPROC**,³⁰ i es va identificar com a versió 1.0. En una segona versió (2.0), aquesta eina ha estat assumida per l'organització objecte de l'estudi, la qual n'ha desenvolupat, amb el mateix model de dades i funcionalitats, una versió evolucionada i hostatjada en un servidor *Microsoft Azure*. Aquesta segona versió és la que conté actualment les dades que han estat objecte d'anàlisi.³¹

²⁸ Bàsicament *Records in Contexts*, del Consell Internacional d'Arxius, *NEDA-MC. Modelo conceptual de descripción archivística*, i *Esquema de Metadatos para la Gestión del Documento Electrónico (e-EMGDE)*, ambdós de l'Administració general de l'estat espanyol.

²⁹ En aquesta guia es desenvolupa especialment la noció de graella o matriu de variables d'un procés, proposada per Antoni de Pàdua Puig-Pey y Pere Guiu en el seu ja clàssic manual (Puig-Pey & Guiu, 2002), i el posterior de 2008 (Puig-Pey & Guiu, 2008).

³⁰ El nom de l'eina, CATPROC, prové de l'acrònim de *Catàleg de Processos*, i es va anomenar així perquè aquesta va ser la primera operació realitzada. Aquest nom s'ha mantingut al llarg de la seva evolució a despit de l'ampliació de l'abast d'aplicació.

³¹ La migració de CATPROC 1.0 a CATPROC 2.0 s'ha realitzat en dues onades, una primera l'any 2020 amb increment de funcionalitats, i la segona el 2022 amb un canvi a nivell d'infraestructura i servidor, com es descriu en el capítol 4.

El desenvolupament de l'eina ha evidenciat progressivament un cert potencial com a sistema de governança transversal de la informació, i també com un instrument eficaç per accelerar els processos de reenginyeria, optimització i dadificació de processos, així com la planificació i el desenvolupament de solucions tecnològiques (preferentment *low code*) per automatitzar aquests processos. Per aquest motiu l'eina ha incorporat millores en la seva interfície d'usuari, i funcionalitats com la possibilitat d'importar fluxgrames i convertir-los automàticament en descripcions estructurades de processos, o la de generar dinàmicament fluxgrames a partir de conjunts d'elements descriptius de processos.

El component analític de la solució desenvolupada són un conjunt de *dashboards* interactius, desenvolupats en *Microsoft Power BI*, i connectats directament a la base de dades de l'eina de recollida. Les anàlisi de conglomerats i correlacions s'han implementat en alguns casos en els propis *dashboards* (activant els components vinculats a R de *Microsoft Power BI*), i en d'altres s'han executat externament utilitzant la solució *Orange Data Mining* versió 3.34.

1.5.2.3 Tria del cas d'aplicació

La metodologia i el disseny conceptual de l'eina de recollida es va validar inicialment amb implementacions puntuals en activitats realitzades en col·laboració amb la Universidad de Costa Rica, amb el Col·legi Oficial de Bibliotecaris i Documentalistes de la Comunitat Valenciana, amb els propis desenvolupadors de la metodologia APS, i també en el context del Màster de Gestió Documental i Informació a les Empreses de la Universitat de Barcelona.³²

Però l'anàlisi plantejat requeria disposar d'una implementació global sobre el conjunt d'una organització, i a ser possible amb la possibilitat d'intervenir activament en la transformació dels seus processos per tal de poder avaluar-ne l'impacte en un entorn bàsicament orientat a dades (*data driven*). El vincle personal amb el Consorci d'Educació de Barcelona, i el moment en el qual es trobava el Consorci, va facilitar aquesta possibilitat. Per aquest motiu, el projecte s'ha basat de forma exclusiva en aquest ens, i s'ha desenvolupat de forma continuada a mesura que s'han pogut anar transformant i dadificant els processos de treball d'aquest ens.

El Consorci d'Educació de Barcelona és l'ens del sector públic català que s'ocupa de la gestió, ordenació i millora de l'educació no universitària a la ciutat de Barcelona. Va ser creat per la Llei 22/1998, de 30 de desembre, de la Carta Municipal de Barcelona, i es va constituir legalment pel Decret 84/2002 del Govern de la Generalitat del 5 de febrer de 2002. El Consorci està format per la Generalitat de Catalunya (en un 60%), a través

³² Prèviament ja s'havien portat a terme algunes experimentacions focals al Departament d'Universitats, Recerca i Societat de la Informació i al Departament de la Vicepresidència de la Generalitat de Catalunya (Serra, 2006; Serra & Casellas, 2007; Serra, 2010; Serra, 2013).

del Departament d'Educació (del qual en fa funcions de Servei Territorial a Barcelona ciutat), i per l'Ajuntament de Barcelona (en un 40%). El seu abast competencial inclou tots els centres educatius de la ciutat de Barcelona (públics i privats, amb concert o sense), i tots els nivells educatius no universitaris: educació infantil, primària, secundària, post-obligatòria, de règim especial (artístics, esportius, etc.), formació d'adults, escoles oficials d'idiomes i totes les variants i recursos de l'educació inclusiva. Anualment gestiona l'educació d'aproximadament 240.000 alumnes (només en el nivell obligatori i post-obligatori) i una plantilla de més de 12.000 professionals de l'educació.

Pel seu caràcter d'ens Consorciat, i per tant d'ens del sector públic (Administració institucional), el Consorci d'Educació de Barcelona gaudeix d'una certa independència funcional, que li permet organitzar-se de forma lleugerament diferent a l'Administració pública territorial. Això no obstant, la seva capacitat tecnològica ve determinada per tres marcs de dependència:

- El Departament d'Educació de la Generalitat de Catalunya pel que fa als sistemes d'informació del sistema educatiu d'abast general (no específics de Barcelona).
- La Generalitat de Catalunya, a través del Centre de Tecnologies de la Informació i les Comunicacions, per al proveïment de serveis i solucions tecnològics, com també de les solucions d'Administració electrònica.
- L'Ajuntament de Barcelona, a través de l'Institut Municipal d'Informàtica, per al proveïment puntual de solucions destinades bàsicament a la gestió del personal municipal.

El marge d'autonomia del Consorci pel que fa a transformació digital queda limitat, per tant a:

- Desenvolupaments propis d'abast limitat.
- Desenvolupaments basats en solucions *low code* al núvol i entorns ofimàtics.
- Transformacions realitzades utilitzant eines transversals d'Administració electrònica (també considerades internament desenvolupaments *low code*).

Pel que fa a la gestió documental, el Consorci, en tant que ens públic gestor de documentació pública, està integrat per normativa en el sistema de gestió documental corporatiu de la Generalitat de Catalunya, i està obligat en conseqüència a utilitzar tots

els instruments tècnics d'aquest sistema.³³ I pel que fa a la gestió de processos, està obligat igualment pel marc legal i instrumental corporatiu.³⁴

La manca de desenvolupament instrumental del Decret 76/2020, de 4 d'agost, d'Administració digital, que regula el govern de la dada a la Generalitat de Catalunya, fa que el Consorci encara no disposi a dia d'avui de directrius ni instruments per fer efectiva la governança de la dada. En aquest àmbit, per tant, el marge d'actuació del Consorci ha estat ampli.³⁵

Més enllà d'aquest marc corporatiu, el Consorci no comptava el 2017, any d'inici del projecte, amb cap sistema de governança de la informació ni res equivalent. [REDACTED]

[REDACTED].³⁶ Aquest punt de partida, combinat amb el limitat però existent marge d'autonomia del Consorci en transformació digital, era el que feia viable iniciar un projecte com el que es plantejava.

1.5.3 Fase 3. Identificació, caracterització i mapatge del substrat informacional

L'any 2017 es va iniciar el projecte. Utilitzant la metodologia establerta inicialment amb la **GMAP**, i l'eina de recollida en la seva primera versió (CATPROC 1.0), es va portar a terme el mapatge de tots els processos del Consorci. Cada procés es va caracteritzar amb els atributs indicats en el model de descripció, juntament amb els relacions a nivell d'agent, però sense entrar encara en el detall de les seves accions ni del substrat informacional (documents i dades).

Un cop construït el mapa de processos, es va elaborar el mapa de sèries documentals, com a derivació natural del mapa de processos, i es va registrar també a CATPROC 1.0. La construcció del mapa de sèries va discórrer en paral·lel a la implantació del

³³ Vegeu la relació completa del conjunt d'instruments Aresta a <https://cultura.gencat.cat/ca/temes/arxiu-i-gestio-documental/gestio-documental/gdgencat/aresta/protocol/instruments/>.

³⁴ Vegeu també <https://administraciopublica.gencat.cat/ca/organitzacio/cataleg-de-procediments/>.

³⁵ Es preveu que en breu començaran a estar disponibles els primers elements normatius i instrumentals que desplegaran el Decret 76/2020.

³⁶ Existien algunes solucions verticals, com la d'ajuts menjador o l'eina de gestió de la despesa, però sota el control exclusiu de la unitat administrativa gestora.

sistema de gestió documental del Consorci, i com a instrument bàsic per a la seva posterior gestió. Un tret característic d'aquest sistema ha estat la substitució del quadre de classificació funcional, un dels instruments del sistema de gestió documental requerit tant per la normativa tècnica de gestió documental (ISO 15489) com pel sistema de gestió documental corporatiu de la Generalitat de Catalunya (SGGDA), per un mapa de sèries documentals amb un mínim component ontològic, el qual contempla quatre nivells de relació entre les sèries documentals.³⁷ Per un motiu de mer compliment normatiu, totes les sèries identificades es van mapar amb el quadre de classificació corporatiu ja definit d'antuvi pel Departament de Cultura i pel Departament d'Educació.³⁸ Sobre cada sèrie es van aplicar les polítiques de conservació i accés ja definides pel model corporatiu, i es van complementar amb polítiques més detallades definides en el marc específic del Consorci.

En paral·lel, es va elaborar el mapa de sistemes d'informació, també a CATPROC 1.0.³⁹ La identificació dels sistemes es va fer mitjançant treball de camp realitzat unitat per unitat, i es va complementar mitjançant accessos de verificació a la xarxa local, fent especial èmfasi en la detecció de la informació estructurada emmagatzemada en solucions d'oficina (ofimàtiques). Aquest inventari es va completar amb la caracterització de cada sistema d'informació, i amb el registre de les interaccions, tant automatitzades com manuals, entre els diferents sistemes.

El darrer pas del mapatge va consistir en l'elaboració del mapa de dades⁴⁰ (o variables d'informació) contingudes en cada sistema, a través de l'accés i anàlisi al model de dades de cada sistema. Això va proporcionar el primer nivell de mapatge de dades, les dades a nivell de sistema d'informació. La construcció d'aquest mapa va aportar un nou element de caracterització de les dades, el de l'objecte descrit, que es va incorporar a partir d'aquest moment a l'anàlisi.

La realitat informacional de qualsevol organització és canviant. Per aquest motiu, aquest conjunt de mapes s'ha continuat mantenint i actualitzant al llarg de tot el desenvolupament del projecte.

³⁷ Atès que un dels elements que el projecte volia posar en qüestió era el model majoritàriament vigent de quadre de classificació funcional d'estructura mono-jeràrquica. Vegeu l'estat de la qüestió per una explicació del motiu d'aquesta transformació.

³⁸ Vegeu novament els instruments tècnics d'Aresta, i concretament el quadre de classificació funcional de la Generalitat a <https://cultura.gencat.cat/ca/temes/arxiu-i-gestio-documental/gestio-documental/gdgencat/aresta/protocol/instruments/qdcfgencat/index.html>.

³⁹ Per a una definició del concepte de sistema d'informació utilitzat en aquesta tesi, vegeu la **GMAP** (capítol 3).

⁴⁰ En aquest punt es va començar a donar forma al concepte de "Dada elemental". Per un major detall sobre el sentit d'aquest concepte vegeu l'estat de la qüestió (capítol 2).

1.5.4 Fase 4. Anàlisi i transformació del substrat informacional

A partir de la fixació del mapa de processos es va iniciar un segon conjunt d'operacions, consistent en analitzar el substrat documental dels processos, i quan ha estat possible transformar-lo per tal d'assolir una execució del procés fonamentalment basada en informació estructurada, el que s'ha anomenat en el marc d'aquesta metodologia dadificació del procés.

Aquest conjunt d'operacions és el que es descriu en la **GDAD**, i es treballa sobre la base d'un procés considerat inicialment de forma individual. S'inicia amb una caracterització detallada del substrat documental, continuant amb la seva optimització a nivell documental, la posterior dadificació a través de la matriu de variables, i la integració final en un model de dades incremental comú a tot el Consorci.

Aquesta seqüència s'ha pogut completar només en aquells casos en què ha estat possible. Durant l'execució del projecte una part dels processos del Consorci s'ha modificat substancialment o de forma continuada, o bé ha deixat de portar-se a terme,



Només en aquells casos en què ha estat possible s'ha pogut completar tant l'anàlisi com la dadificació del procés.⁴¹

Aquesta dadificació, per poder-se fer efectiva, ha requerit en molts casos el desenvolupament de solucions tecnològiques, majoritàriament amb tecnologia *low code*, per substituir habitualment una solució ofimàtica preexistent. Aquest fet ha incrementat l'esforç a nivell de procés i ha prolongat el temps necessari per a poder disposar d'una base per a l'anàlisi d'aquell procés.

El principal producte d'aquesta anàlisi ha estat la construcció progressiva d'un segon mapa de dades, el mapa de dades a nivell d'evidència de procés, en el qual s'ha incorporat un element ja apuntat en la metodologia APS: el caràcter autòcton o de ròssec de les variables d'informació (o aparicions/usos de cada dada elemental). Aquest ha esdevingut, per als processos on s'ha pogut desenvolupar, un dels elements primordials per a l'anàlisi de relacions entre les entitats identificades en la fase anterior.

⁴¹ Per aquest motiu, l'anàlisi que es porta a terme en els capítols 6 i 7 es basa en una sèrie de mostres seleccionades en funció del nivell de dadificació i descripció que ha estat possible assolir en cada cas.

1.5.5 Fase 5. Anàlisi de les dades i extracció de conclusions

A partir de la informació recollida a CATPROC, ja en versió 2.0, i en resposta als objectius plantejats, s'ha realitzat una doble anàlisi. En primer lloc, una anàlisi preferentment quantitativa orientada a confirmar o desestimar les hipòtesis plantejades, i intentar extreure possibles patrons d'identificació i tractament. I en segon lloc, una anàlisi qualitativa orientada a validar la operativitat de la solució metodològica i tecnològica utilitzada per a la recollida de la informació, i també, des d'un punt de vista estratègic, la seva aplicabilitat a d'altres organitzacions com a sistema integral de govern de la informació, tant pel que fa al model de dades i sistema d'identificació i transformació, com pel que fa als indicadors de seguiment i avaluació de la qualitat i característiques de la informació.

1.5.5.1 Anàlisi quantitativa per als objectius teòrics

L'anàlisi orientada a confirmar o desestimar les hipòtesis s'ha desenvolupat a tres nivells: descripció al nivell d'entitat, anàlisi al nivell d'entitat, i anàlisi comparativa final.

La descripció s'ha realitzat per a totes les entitats analitzades (processos, sèries documentals, sistemes d'informació i dades elementals). De cada entitat s'ha generat un conjunt d'indicadors descriptius que inclou:

- Indicadors descriptius simples (agregats) de l'entitat analitzada, basats exclusivament en atributs de l'entitat.
- Indicadors descriptius derivats de les relacions d'aquesta entitat amb la resta d'entitats.
- Indicadors sobre les dades manifestades en aquella entitat.
- Una representació gràfica i interactiva de les relacions entre les instàncies d'aquella entitat (el mapa de relacions de l'entitat).

L'anàlisi al nivell d'entitat s'ha realitzat també per a totes les entitats analitzades (processos, sèries documentals, sistemes d'informació i dades elementals). Per a cada entitat s'ha generat una sèrie d'indicadors, que inclouen:

- Indicadors i mesures calculats a partir dels indicadors descriptius.
- Classificacions i assignacions de categories a partir dels indicadors anteriors.

L'anàlisi comparativa final ha combinat els indicadors anteriors per intentar descobrir patrons, tant en les pròpies entitats com en les relacions entre entitats. Per a fer-ho s'ha utilitzat en alguns casos dues tècniques estadístiques: l'anàlisi de correlació entre

determinades variables considerades potencialment significatives, i l'anàlisi de conglomerats per detectar patrons i possibles relacions entre aquests patrons, com també per a seleccionar mostres d'estudi.

Sobre aquesta base quantitativa s'han formulat les conclusions que s'exposen en el capítol 8.

1.5.5.2 Anàlisi qualitativa per als objectius instrumentals

L'anàlisi qualitativa s'ha realitzat a partir de l'aplicació pràctica dels components desenvolupats (guies, model de dades, CATPROC I *dashboards*) en dues línies d'accions.

Pel que fa a la validació a un nivell operatiu, s'ha avaluat la validesa del model a partir de les següents accions:

- Valorar l'esforç de transformació d'un tràmit fins a la disponibilitat en producció de la seva forma dadificada.
- Valorar l'encert en les identificacions de les instàncies de les diferents entitats.
- Valorar la facilitat amb la qual equips externs podien utilitzar els conceptes del model i realitzar identificacions correctes.
- Valorar si ha estat possible mantenir al dia i actualitzats els mapes en un context real de canvi continu a nivell de normativa, processos, sistemes d'informació i demandes de dades.
- Valorar la utilitat del mapa de sistemes d'informació i del mapa de dades en funció de la seva utilització continuada per part de l'Oficina de Dades del Consorci.

Pel que fa a la validació al nivell estratègic, l'avaluació s'ha portat a terme amb les següents accions:

- Generació de mapes i visualitzacions quantitatives per a la presa de decisions sobre la modificació de processos i la prioritització d'accions de millora.
- Generació de visualitzacions per a la planificació del desenvolupament de nous sistemes d'informació o la transformació, evolució o integració de sistemes ja existents.

Amb el resultat d'aquesta avaluació s'ha incorporat millores a les dues guies metodològiques i al conjunt d'indicadors proposat inicialment, per tal de proposar un model que pugui ser exportable a d'altres organitzacions de característiques similars.

1.6 Estructura de la tesi

1.6.1 Estructura dels capítols

Aquesta introducció constitueix el capítol 1 de la tesi. El capítol 2 fa un viatge per l'estat de la qüestió dels conceptes que entren en joc en el sistema de govern de la informació que posteriorment s'analitza. S'hi repassen, des del punt de vista de la disciplina arxivística i de la gestió documental, l'evolució recent dels conceptes de document i d'agrupació documental, especialment centrat en el concepte de sèrie documental, com també del context que justifica la utilització funcional d'aquests conceptes (procés, tasca, agent productor, etc.). També s'inclou en aquest estat de la qüestió l'anàlisi dels conceptes de sistema d'informació o magatzem de dades, i de la pròpia dada, incorporant visions pròpies de les disciplines de les tecnologies de la informació i de l'analítica de dades.

El capítols 3, 4 i 5 presenten la metodologia utilitzada per a la recollida de la informació. L'eix d'aquesta metodologia són dues guies metodològiques que s'han utilitzat com a marc conceptual i procedimental per a la recollida de la informació del projecte, i que s'expliquen en el capítol 3. La primera guia presenta un mètode per a la identificació i mapatge de processos, sèries, sistemes i dades en el procés de construcció d'un sistema de govern de la informació. I la segona guia presenta una metodologia per a la dadificació de processos a través de l'anàlisi de les seves evidències d'informació, i per al posterior desplegament d'aquestes dades en la construcció d'un model de dades global de l'organització.

Per a la recollida de la informació resultant de l'aplicació d'ambdues guies s'ha desenvolupat una aplicació informàtica. En el capítol 4 es descriu aquesta aplicació, les seves funcionalitats, i es detalla el model de dades en el qual es basa. Aquest model de dades sintetitza l'aportació conceptual que rau en el fons de la metodologia proposada.

La metodologia es completa amb un conjunt d'indicadors per a la mesura, avaluació i millora dels actius d'informació d'una organització en el marc d'un sistema de governança de la informació. Aquests indicadors es descriuen en el capítol 5.

Els capítols 6 i 7 presenten els resultats de l'aplicació d'aquesta metodologia sobre el cas del Consorci d'Educació de Barcelona. Aquesta aplicació es descriu detalladament en el capítol 6, utilitzant els indicadors que la metodologia proposa en el capítol 5, i es detallen les anàlisi realitzades sobre el conjunt de les dades recollides, comentant per cada indicador el resultat obtingut. A continuació, en el capítol 7 es posa el focus en els objectius generals de la tesi, i s'utilitzen les dades del Consorci per analitzar les relacions entre les diferents entitats i intentar identificar-ne patrons que serveixin d'evidència per a les hipòtesis plantejades.

La resposta a aquestes hipòtesis, ja amb caràcter general, es recull en el capítol 8, on s'exposen les conclusions derivades de l'anàlisi anterior, i es formulen algunes propostes aplicables al disseny de sistemes de govern de la informació, com també la valoració sobre el possible ús de la metodologia utilitzada en contextos similars.

Finalment, el capítol 9 conté la bibliografia utilitzada, amb excepció de les normes de descripció arxivística, que se citen en l'apartat 4.5 del capítol 5, i de les pàgines Web institucionals, que se citen en tots els casos a peu de pàgina.

1.6.2 Versió per a publicació

La Llei 19/2014, del 29 de desembre, de transparència, accés a la informació pública i bon govern, en el seu article 21.1.a), estableix com un dels límits al dret d'accés a la informació pública, i en conseqüència a la possibilitat de fer pública informació d'un ens públic, el fet que el coneixement o la divulgació de la informació pugui comportar un perjudici per a la seguretat pública.

Atès que la tesi conté informació detallada sobre les característiques dels sistemes d'informació del Consorci d'Educació de Barcelona, la qual es considera sensible per motius de seguretat informàtica, s'ha elaborat una versió de la tesi, orientada a la publicació en obert, en la qual s'ha ocultat parcialment informació dels capítols 1, 6 i 7.

La versió que s'ofereix amb aquest document correspon a la versió amb ocultació parcial de la informació considerada sensible per al Consorci d'Educació de Barcelona. S'ha conservat en el document la referència als apartats ocultats per a mantenir la intel·ligibilitat de l'estructura. Les parts del document que s'han ocultat són les següents:

- Paràgraf a l'apartat 1.5.2.3 *Execució de les accions de recollida*
- Paràgraf a l'apartat 1.5.4 *Fase 4. Anàlisi i transformació del substrat informacional*
- Apartat 6.6.1 *Mètriques descriptives dels sistemes d'informació*
- Taula 6-49: *Entorns tecnològics de les sèries documentals del grup focal*
- Taula 6-62: *Relació dels deu sistemes d'informació amb un índex de connectivitat més alt*
- Taula 6-69: *Classificació de seguretat dels sistemes d'informació per abast*
- Apartat 6.6.5.3 *Classificació dels sistemes pel nivell de seguretat de les dades*
- Apartats 7.1.1 a 7.1.8 de l'apartat 7.1 *Anàlisi de les característiques de les sèries documentals dadificades del Consorci d'Educació de Barcelona*
- Taula 7-34: *Nivell de seguretat dels sistemes d'informació del grup focal*

- Paràgraf a l'apartat 7.3 *Objectiu 2: característiques de l'ecosistema tecnològic de la sèrie documental dadificada*
- Taula 7-38: *Sistemes d'informació amb més processos vinculats*
- Taula 7-42: *Caràcter estimat dels sistemes d'informació vinculats a més processos*
- Taula 7-45: *Models tecnològics dels sistemes d'informació del primer clúster*
- Taula 7-46: *Models tecnològics dels sistemes d'informació del segon clúster*
- Taula 7-47: *Models tecnològics dels sistemes d'informació del tercer clúster*
- Taula 7-48: *Sistemes d'informació del tercer clúster que gestionen dades mestres*
- Taula 7-49: *Sistemes d'informació amb un major nombre de coincidències de dades amb altres sistemes*
- Taula 7-50: *Sistemes d'informació amb un índex de coincidència major*

1.7 Guia de lectura

A continuació es proposen tres itineraris de lectura de la tesi, en funció de l'objectiu de preferència:

- A. Per conèixer les **aportacions conceptuals i conclusions** d'aquesta tesi, es recomana llegir en primer lloc la introducció (capítol 1), a continuació els apartats inicials de les dues parts de l'estat de la qüestió (apartats 2.1, 2.2.1, 2.3.1 i 2.3.6 del capítol 2), i finalment el capítol dedicat a les conclusions (capítol 8). La fonamentació quantitativa de les conclusions es pot consultar en el capítol 7 a partir de l'apartat 7.2.

- B. Per conèixer la **metodologia proposada**, es recomana llegir en primer lloc la segona part de l'estat de la qüestió (apartats 2.1 i 2.3 del capítol 2), i després els capítols dedicats a les dues guies metodològiques (capítol 3), el model de dades (capítol 4) i els indicadors de governança (capítol 5). La valoració de l'aplicació de la metodologia es pot consultar a l'apartat 8.2 del capítol 8.

- C. Per conèixer el **cas d'aplicació al Consorci d'Educació de Barcelona**, es recomana llegir els capítols 6 i 7, prèvia consulta de l'apartat 1.5.2.3 del capítol 1, de les dues guies metodològiques (capítol 3) i dels indicadors de governança (capítol 5). Les conclusions (capítol 8) esdevenen en aquest cas una abstracció de les evidències del cas d'aplicació del Consorci.

Capítol 2. Estat de la qüestió

2.1 Plantejament i estructura de l'estat de la qüestió

La gestió de la informació a les organitzacions⁴², en la forma en què la coneixem actualment, es va configurar després de la Segona Guerra Mundial, i va comportar l'adopció regular de la gestió documental per part dels entorns empresarials (McDonald, 1995; Svärd, 2014). En l'àmbit anglosaxó es va començar a configurar la doctrina de la gestió de documents (*records management*) de la mà de Theodore R. Schellenberg (Schellenberg, 1956, 1965). I en paral·lel, l'eclosió de la tecnologia informàtica va configurar progressivament un nou àmbit de coneixement al voltant de la gestió de la informació basada en dades (McDonald, 1989, 1995).

A mesura que ocupaven més espai dins l'estructura de les organitzacions, el diàleg entre ambdós àmbits fou cada cop més complex (Bearman, 1991; McLeod, 2008; Yeo, 2011). Els sistemes d'informació eren inicialment sistemes tancats, on la informació en forma de dades era captiva del sistema, i no se'n podia plantejar cap gestió diferenciada (McDonald, 1995; Julbe, 2023). No fou fins que la tecnologia informàtica va evolucionar cap a una estructura client-servidor i un desenvolupament en tres capes de les aplicacions, que les dades, segregades de l'aplicació i reutilitzades per diferents aplicacions, es van poder considerar un actiu d'informació independent, i van poder ser objecte d'una gestió diferenciada (Julbe, 2023).

Contemporània a aquesta evolució és la transformació massiva dels documents en paper a documents electrònics, i la possibilitat de gestionar-los dins d'aplicacions informàtiques especialment orientades a la gestió d'objectes (els gestors documentals i les bases de dades orientades a objectes) (*Committee on Electronic Records*, 1997; Jansen, 2015). La possibilitat de gestionar descripcions vinculades intrínsecament als objectes descrits en forma de metadades és el primer vincle clar entre ambdós àmbits (Frendo, 2007; Niu, 2013). A partir d'aquí, l'aparició dels metallenguatges i l'evolució en la codificació dels documents electrònics ha acabat fusionant la gestió d'objectes amb la gestió de dades estructurades i/o estructurables.

En el moment en què les organitzacions disposen de la capacitat tècnica per gestionar de forma coordinada la informació estructurada i no estructurada, comença a apuntar-se la possibilitat d'un govern integral de la informació (Kooper et al., 2011; Kampffmeyer,

⁴² Quan parlem d'informació ens referim a la informació de caràcter administratiu, aquella generada per a usos concrets i de caràcter evidencial, i no a la informació adreçada a generar coneixement, pròpia de la biblioteconomia i la documentació.

2023), que aglutina les actuacions purament tecnològiques, les actuacions de tractament i organització de la informació, i la vetlla per la seguretat i pel seu valor evidencial (*compliance*). Aquesta idea de govern integral, malgrat que formulada de forma general i des de diferents àmbits (Harries, 2009; Loshin, 2010; Kooper et al., 2011), ha inspirat l'evolució metodològica d'aquests àmbits, i especialment de la gestió documental (Brooks, 2019; Upward, 2019). El desenvolupament de la família de normes ISO 30300 cal ubicar-la també en aquest context (Bustelo, 2011).

La darrera gran evolució en l'àmbit de la gestió de la informació ha vingut de la mà de l'explotació de la informació (Mayer-Schönberger & Cukier, 2014). L'augment de la capacitat de processament massiu de dades, producte de la conjunció de l'increment de la potència de processament, de la capacitat d'emmagatzematge, i de la capacitat de captura automatitzada de dades, ha desplaçat el focus d'interès principal cap a l'explotació de la informació i el seu ús en processos automatitzats autònoms (Lappin, 2020a). Aquest fet ha provocat l'emergència d'un tercer àmbit, el de la governança de les dades (Khatri & Brown, 2010; Loshin, 2015; Ritter et al., 2015), el qual, a diferència del govern integrat de la informació, no expressa una vocació clara d'integració amb el conjunt de la gestió de la informació (Earley & Henderson, 2017).⁴³ L'impacte sobre la gestió documental no ha estat fins el moment positiu, i en tot cas tampoc s'han obert vies de col·laboració formals (Loshin, 2010; Benfeldt et al., 2020).

L'objectiu d'aquesta tesi és copsar quin paper ha de jugar la sèrie documental, i per extensió els diferents nivells d'agrupacions documentals arxivístiques, pròpies també de la gestió documental, en els ecosistemes informacionals actuals, altament automatitzats i basats fonamentalment en dades. I a partir d'aquesta comprensió, valorar si el concepte de sèrie documental segueix essent útil per a l'aplicació de les polítiques de gestió documental (bàsicament conjunts de regles de conservació, accés i ús).

Per aquest motiu, el primer que cal fer és conèixer no només l'evolució del concepte de sèrie documental, sinó també de l'ecosistema informacional en el qual aquest concepte ha viscut. Aquest ecosistema inclou allò que és objecte de gestió (els documents i les dades), allò que en motiva la gestió (els processos i activitats), i els instruments que s'utilitzen per aquesta gestió (els sistemes d'informació i aplicacions). L'abast és tant ampli, i ha estat tractat des de tantes disciplines diferents, que aquest estat de la qüestió s'ha hagut de limitar, per qüestió de factibilitat, a un determinat conjunt de fonts: aquelles vinculades directa o tangencialment amb l'arxivística i la gestió documental, o bé les més citades per la bibliografia de referència de cada àmbit.⁴⁴

⁴³ El *DMBOK2* de *DAMA International*, actualment el model de referència per al disseny de sistemes de governança de les dades, contempla com un dels seus onze àmbits operacionals a la gestió documental, però bàsicament per a gestionar la documentació derivada dels projectes de gestió de dades.

⁴⁴ Llevat de les fonts relatives a sistemes d'informació i a la ciència de dades, que han estat incorporades únicament per aquells aspectes que poden tenir relació amb la identificació de sèries documentals: els conjunts de dades i els entorns que les produeixen i emmagatzemen.

Així mateix, s'ha prioritzat en aquest estat de la qüestió la visió que es vincula al mandat implícit en la disciplina arxivística: el d'una gestió eficaç que garanteixi l'autenticitat, la integritat, la conservació i el bon ús de la informació, especialment per part i en el marc de l'organització productora. No s'ha inclòs, per tant, l'aportació que la disciplina que gestiona cada objecte fa tant a nivell metodològic com conceptual, llevat d'aquells aspectes relacionats directament amb els objectius de la gestió documental.

L'estat de la qüestió s'ha estructurat en dos grans blocs. En el primer bloc s'analitza de quina forma s'està gestionant la informació en el nou marc que configura la ciència de dades, i la tendència a orientar la gestió de les organitzacions cap a una gestió basada en dades (*data driven*). Per aquesta anàlisi se cerquen els punts de contacte entre els diferents productes basats en dades i el concepte tradicional de document, explorant la dicotomia entre sistemes de gestió datacèntrics i docucèntrics. En aquesta anàlisi es presta especial atenció als entorns on es generen les dades (els sistemes d'informació i/o aplicacions), i aquells on s'emmagatzemen (diferents tipus de bases de dades o dipòsits de dades).

I en el segon bloc s'analitza quina ha estat l'evolució del concepte de sèrie documental, i per extensió d'agrupació documental arxivística, a partir de la configuració de la gestió documental moderna.⁴⁵ S'analitzen les diferents concepcions de sèrie documental, i especialment el seu origen com a derivació documental d'un conjunt de funcions, activitats o processos. Es presta una atenció especial a aquest vincle, en tant que determinant de la significació última del concepte de sèrie documental, i també de la configuració dels instruments de classificació i descripció que han conformat l'arsenal operatiu dels sistemes de gestió documental.

L'anàlisi també ha servit de base per a la construcció de la metodologia utilitzada per a la recollida i explotació de la informació sobre el cas pràctic objecte d'estudi. I és en aquest punt on entronca amb el segon objectiu de la tesi: donar forma a un model metodològic vàlid per a la governança de la informació en una organització de dimensions i característiques similars al Consorci d'Educació de Barcelona. Amb aquesta finalitat s'han analitzat els models de descripció més utilitzats per la gestió documental i l'arxivística, i quin model conceptual rau en cadascun d'ells, amb especial atenció a les relacions que estableixen entre les diferents entitats. Per cadascuna de les entitats o conceptes que participen en la gestió de la informació, s'ha analitzat com ha estat conceptualitzada en els diferents models, quins elements de metadades s'han utilitzat per a descriure-la, i per quina formulació del concepte optem en el context d'aquesta tesi. Aquesta formulació final és la que s'ha utilitzat després per a la definició del model de dades emprat per recollir la informació del cas d'aplicació, per a la construcció de les dues guies metodològiques de mapatge i dadificació utilitzades també en la recollida d'informació, i per al conjunt d'indicadors i mesures proposats per a

⁴⁵ Entenen moderna com la gestió documental que ha actuat des del principi del cicle de vida i que ha incorporat de forma natural el document electrònic. Agafem doncs com a punt de partida el model de gestió documental recollit en la norma ISO 15489.

l'anàlisi i optimització del mapa informacional. Aquesta part de l'estat de la qüestió es troba recollida en part en aquest capítol, i en part en les introduccions de les dues guies metodològiques (apartat 3.1.2 capítol 3) i del model de dades (apartat 4.5 capítol 4).

2.2 El constructe tecnològic: documents, dades i sistemes d'informació

2.2.1 La dialèctica document-dada i la crisi del concepte de document

Qualsevol actuació d'una organització genera dades (Malmsten, 2013; Smit et al., 2017). Documentar una activitat implica recollir totes les dades que generen cadascuna de les actuacions, accions o, utilitzant terminologia anglosaxona, *transaccions* que la componen.⁴⁶ Si aquestes transaccions documentades reben la forma d'un objecte independent i unitari, és a dir un document, o pel contrari es mantenen com un simple conjunt de dades, és un fet circumstancial i transparent per als protagonistes de l'activitat (Bak, 2012). Malgrat les seves característiques tecnològiques siguin totalment diferents, la seva vinculació amb l'activitat original es mantindrà intacta, i amb ella el seu valor probatori o evidencial (Bak, 2016b).

Aquesta idea apunta a l'essència de la dicotomia amb la qual obrim aquest estat de la qüestió. La distinció entre documents i dades es formalitza inicialment amb la distinció entre informació estructurada i no estructurada (McDonald, 1989), considerant els documents com informació no estructurada, idea qüestionada per Geoffrey Yeo (Yeo, 2011). Malgrat la imprecisió d'aquesta distinció, el seu arrelament, tant en l'arxivística com en la ciència de dades, es manté plenament vigent (Diedel, 2020; Makhoulf Shabou et al., 2020; Lo Giudice et al., 2019).

El debat entre informació estructurada i no estructurada ha estat sempre contaminat de confusions (Borglund & Engvall, 2014; Yamaoka & Gauthier, 2013). A mitjans dels anys 90 l'arxivística es plantejava preguntes sobre si les bases de dades eren documents, o contenien documents, o si una taula de base de dades era un document. Així van sorgir els conceptes de document-objecte (*document like object*, o DLO), equivalent a informació no estructurada, i document-sistema (*document like system*, o DLS), equivalent a informació estructurada.⁴⁷ Es considerava document-objecte totes aquelles unitats d'informació que podien ser gestionades de forma autònoma, dins un sistema de

⁴⁶ Per a fer referència a la seqüència d'activitats que porta a terme una organització per assolir un determinat objectiu podem parlar d'actuació, procediment o tramitació. En aquest estudi utilitzaré puntualment el terme *tramitació* perquè és el més utilitzat quan es fa referència a les activitats que es realitzen en línia o de forma telemàtica (*tramitació electrònica*). S'entendrà doncs tramitació administrativa com aquella seqüència d'activitats que porta a terme una organització, pública o privada, per assolir un objectiu concret. Habitualment segueix un procediment previsible o reglat, tot i que no és preceptiu que ho faci. L'estudi se centrarà preferentment en les tramitacions administratives, tant reglades com no reglades, però també inclourà les transaccions o tràmits que es realitzen de forma descontextualitzada (p.e. les comunicacions protocol·làries).

⁴⁷ Sobre els conceptes de DLO i DLS vegeu l'apartat 1.1.1 del capítol 1.

fitxers, en forma de fitxers o conjunts de fitxers. I es considerava document-sistema totes aquelles unitats d'informació captives d'una estructura tecnològica, fora de la qual no podien tenir existència en tant que unitats d'informació.

Aquesta dicotomia dades soltes (i estructurades en el seu context de tramitació), o dades agrupades en objectes independents formant documents-objecte, va donar origen a que es parlés de dos models per documentar la realitat de les organitzacions: el model datacèntric, que prioritzava la gestió de dades estructurades, i el model docucèntric, centrat en la gestió de col·leccions o agrupacions d'objectes discrets (Harries, 2000). El model docucèntric era el més proper conceptualment a la disciplina arxivística, mentre que el model datacèntric era l'entorn preferent per als tecnòlegs. Durant un temps aquesta dicotomia va determinar les orientacions preferents d'ambdós col·lectius en la transformació digital de les activitats, i especialment la preferència dels arxivers per eines de gestió de documents electrònics (EDMS).⁴⁸

L'evolució recent de les formes tecnològiques que prenen les unitats d'informació està esborrant progressivament aquesta diferència entre informació estructurada i no estructurada. De fet, John McDonald ja va intuir l'any 1988⁴⁹ que el punt de contacte entre ambdues dimensions es trobaria en l'ofimàtica col·laborativa, i que actuar sobre aquest àmbit seria clau per aconseguir que la gestió documental fos realment exhaustiva (McDonald, 1989). Els canvis en la semàntica dels formats dels documents electrònics, dels documents-objecte, així com la seva progressiva estandardització, ha fet que es puguin processar com informació estructurada. El fet de poder-los tractar i explotar gairebé de la mateixa manera que la informació de les bases de dades, gràcies sobretot a l'aparició de les bases de dades NoSQL, ha fomentat encara més aquesta consideració. Per aquest motiu alguns autors afirmen que el document, l'element que dona nom a la gestió documental, i per extensió a l'arxivística, està en dissolució: en un procés dadificat, el document, tal com l'entén i el descriu l'arxivística, tendeix a desaparèixer (Yeo, 2007, 2011; Smit et al., 2017). Segons Chris Walker (Walker, 2023):

“Whether the information is structured, semi-structured, or unstructured (there`s no such thing) makes no difference. [...] The only thing that matters is whether or not the information is needed by the organization to either conduct business or meet obligations.”

Però per analitzar adequadament el vincle del document amb la dada, voldríem abans aprofundir un moment en la definició arxivística del concepte de document.

El concepte de document en l'àmbit arxivístic ha estat dotat des dels seus inicis d'un component patrimonial, d'un component corpori. La definició tradicional fa present aquesta corporeïtat en el concepte de “suport”, quan es defineix el document com

⁴⁸ Per una explicació dels EDMS vegeu l'apartat 2.2.3.6 d'aquest capítol.

⁴⁹ Any de la conferència recollida en l'article de McDonald 1989.

“informació fixada en un suport” (Buckland, 1991). Una identificació que actualment es manté vigent, tant en la normativa tècnica com jurídica.⁵⁰

L'altre concepte que apareix en la definició esmentada, el de “informació”, simbolitza aquesta altra dimensió del document: la no corpòria, la significativa. El concepte “informació” té una dimensió abstracta, però adquireix existència operativa (és a dir, pot ser objecte de gestió) quan ens hi referim com a una unitat d'informació (Dahlberg, 1993). Michael Buckland anomenava a aquesta dimensió abstracta la “informació com a procés”, mentre que la dimensió operativa corresponia a la “informació com a cosa” (Buckland, 1991). En la seva forma més simple, podem interpretar que construïm una unitat d'informació a partir de la suma de dades, i que aquesta unitat es converteix en un significat quan es combina amb la capacitat d'interpretació (el coneixement previ) del receptor, idea subjacent en l'equació infològica de Langefors (Langefors & Dahlbom, 1995). Aquesta suma de dades construeix un missatge, un significat, que és el que realment pot ser utilitzat, a vegades de forma absoluta i d'altres en un context determinat, i que en l'àmbit de la gestió documental, quan se'l dota de persistència en el temps, s'anomena document. Malauradament, els conceptes dada i informació s'han considerat sovint, en moltes disciplines, com a sinònims (Borglund & Engvall, 2014; Lueders, 2021), sense unes fronteres clares que els defineixin, alimentant una certa “inflació terminològica” (Sundgren, 2003), i allunyant la gestió de la informació estructurada del concepte de document i de la gestió documental (Brooks, 2019).

El projecte PREMIS de preservació digital⁵¹ feu una aportació molt rellevant a aquesta conceptualització del document. La significació a la qual hem fet esment correspon al que PREMIS anomena “entitat intel·lectual”, la qual s'instancia en combinacions d'objectes (fitxers) o cadenes de bits formant diferents “representacions” (Caplan, 2021).⁵² Aquesta distinció entre entitat intel·lectual, és a dir un significat, i la seva concreció física temporal, és de gran importància per a la gestió de la documentació

⁵⁰ A títol d'exemple, la normativa catalana defuig donar una definició de document, i remet a la definició patrimonial que s'identifica amb les dimensions esmentades per Buckland. Així doncs, el *Decret 76/2020, de 4 d'agost, d'Administració digital*, en l'article 22, remet a la *Llei 10/2001, de 13 de juliol, d'arxius i documents*, la qual a la seva vegada, en l'article 2.a), remet a la *Llei 9/1993, del 30 de setembre, del patrimoni cultural català*, que en l'article 19.1 acaba definint document com “[...] tota expressió en llenguatge oral, escrit, d'imatges o de sons, natural o codificat, recollida en qualsevol mena de suport material, i qualsevol altra expressió gràfica que constitueixi un testimoni de les funcions i les activitats socials de l'home i dels grups humans, amb exclusió de les obres d'investigació o de creació.” I també en l'article 22, en definir el concepte d'actiu digital, el Decret 76/2020 diu que “1. A l'efecte d'aquest Decret, i d'acord amb la definició de document que estableix la normativa vigent, l'actiu digital és la dada o la informació obtinguda o generada en el desenvolupament de les funcions i competències dels subjectes de l'article 2 d'aquest Decret que requereix ser documentada i conservada durant un període de temps determinat. Concretament, són actius digitals els fitxers de dades, d'imatges, de sons, de text i multimèdia, així com qualsevol tipus de dada o de representació d'informació, entre els quals hi ha els documents administratius.”

⁵¹ Vegeu <https://www.loc.gov/standards/premis/>.

⁵² Més tard els Arxius Nacionals del Regne Unit en digueren “manifestacions” (Brown, 2014).

electrònica, i especialment per a la gestió d'aquells documents on el canvi de forma externa no n'altera el significat, com són els documents basats exclusivament en dades (Duranti et al., 2002). Miguel Ferreira va desgranar aquestes dues dimensions en quatre, distingint, dins la implementació, entre la dimensió física i la dimensió lògica (o d'interpretació), i dins la significació entre la dimensió conceptual (el significat pur) i la dimensió experimentada (el significat percebut, en la línia de l'equació infològica de Langefors (Ferreira, 2006). Però el caràcter *representacional* del document no està exempt de controvèrsia, atès que depèn d'una interpretació humana que, malgrat el document es mantingui immutable, pot patir canvis en el temps (Yeo, 2012).

És a partir d'aquesta dimensió conceptual o de significat que el document es pot desagregar en unitats més petites (dades) i agregar en unitats de major abast (agrupacions documentals amb significació pròpia), convertint el document, en tant que "entitat intel·lectual", en un agregat temporal d'unitats atòmiques d'informació, les dades, que permet construir i comunicar un significat. El concepte d'unitat d'informació, entès com entitat intel·lectual instanciada, converteix les dades en un ens gestionable, de forma autònoma o dins un context conceptual i tecnològic determinat, atès que les dota d'uns límits reconeixibles. Aquesta condició objectual de la unitat d'informació és el que l'associa al concepte d'ús (o usos) de la informació, i és en funció d'aquests usos que pot generar nous conceptes (Becker, 2018).

Quan una unitat d'informació és coneguda per una persona, es converteix per aquella persona en el que s'anomena coneixement (Davenport & Prusak, 1998), el qual esdevé l'ús primer i més elemental de les unitats d'informació. Però no és un ús necessàriament productiu, és a dir que produeixi quelcom assimilable a un bé o servei. Per tant, és un ús sense un valor intrínsec. En el moment en què fixem la mirada en el valor d'aquest coneixement, és a dir en el seu caràcter de bé o producte (cosa amb valor monetari o d'intercanvi), passem a definir-ho com a actiu d'informació (*information asset*). És així com ho descriu el *Decret 76/2020, de 4 d'agost, d'Administració digital*, que regula la nova onada de desplegament de l'Administració electrònica a la Generalitat de Catalunya.

Però quan fixem la mirada en el caràcter de la unitat d'informació entesa com a servei (cosa que pot produir efectes, és a dir una acció en potència), no apliquem la definició passiva d'actiu o recurs d'informació, sinó que parlarem de document administratiu o d'arxiu (Franks, 2013). En aquest cas, els efectes que pugui produir la unitat d'informació venen condicionats per una variable: el significat, o entitat intel·lectual que representa la unitat d'informació ha de ser cert, o en tot cas ha de tenir presumpció de veracitat (Kastenhofer, 2015; Bak, 2016b). És a dir, hem de poder pensar que les dades que construeixen aquell significat són certes. Sense aquesta condició de veracitat

presumpta, la unitat d'informació no podrà tenir mai el caràcter de document administratiu.⁵³

Aquesta presumpció de veracitat de l'evidència prové de la capacitat tècnica de verificar tres condicions de la unitat d'informació (Duranti & MacNeil, 1996; Soler, 2011; Rogers, 2015):

- **L'autenticitat:** la possibilitat de verificar que el context de creació (o d'agregació) de la unitat d'informació, expressat en la pròpia unitat d'informació o en les seves relacions, és cert.
- **La integritat:** la possibilitat de verificar que cap dels elements o dades que componen la unitat d'informació ha estat modificat ni n'ha estat separat des del moment en què es va registrar el context de creació.
- **La fiabilitat:** la possibilitat, a partir del coneixement del context de creació, d'assignar un determinat nivell de confiança a la informació de contingut de la unitat d'informació.

Aquest èmfasi en la presumpció de veracitat ha propiciat noves denominacions del document, entre elles la de "evidència electrònica". El concepte d'evidència electrònica prové de l'àmbit jurídic.⁵⁴ Però el concepte d'evidència electrònica, més enllà de la seva lectura jurídica, té un important potencial d'utilització en l'àmbit arxivístic. El concepte d'evidència electrònica, documental o no, és el que millor s'ajusta al concepte anglosaxó de *record* (o registre evidencial d'una transacció). Com diu Alejandro Delgado, la traducció del mot *record* per "document administratiu", la més utilitzada fins ara al nostre entorn,⁵⁵ és una traducció forçada, ja que no recull la complexitat d'aquelles formes d'enregistrar la informació administrativa-evidencial-transaccional que no s'ajusten al concepte de document-objecte (Delgado, 2007).⁵⁶

⁵³ Per aquest motiu, enlloc d'utilitzar el concepte de document administratiu, que provoca força discussió en la literatura científica i tècnica, he preferit utilitzar el concepte d'evidència, és a dir unitat d'informació presumptament certa.

⁵⁴ El concepte d'evidència electrònica s'ha d'ubicar en un context de forta dependència d'un ordenament jurídic de tradició llatina, basat en la norma escrita i en l'acreditació sistemàtica de l'autenticitat en totes les actuacions. L'ús de preceptes propis del dret de tradició anglosaxona en els projectes d'administració electrònica ha estat al nostre país un punt de conflicte constant entre juristes i tecnòlegs, i ha condicionat fortament l'èxit d'alguns d'aquests projectes (Delgado, 2007). Cal tenir present que els models de referència, tant per a l'administració i el comerç electrònic com per a la preservació digital, provenen gairebé en exclusiva del món anglosaxó, i que problemes de caire instrumental que en el nostre entorn immediat porten al col·lapse d'un projecte no són vistos com a tals en els àmbits on s'originen les solucions (Soler, 2011).

⁵⁵ És la que es va utilitzar en la traducció espanyola de la norma ISO 15489.

⁵⁶ Per aquest motiu sovint s'opta per la traducció literal "registre", com és el cas de la traducció espanyola de la norma ISO 9000. Un record pot ser un document administratiu, un conjunt o seqüència de documents administratius, o un document recapitulatiu, el que l'aproxima al nostre concepte de registre, i també a la idea de fitxer o base de dades. De fet, la norma MoReq2 diu clarament que un *record* pot estar compostat d'un o més documents.

Aquest concepte de document administratiu o d'arxiu ha estat un permanent focus de debat, atesa la diferent nomenclatura, i els matisos de concepte, que rep a les diferents tradicions arxivístiques, i que queden reflectides en la normativa de descripció arxivística (Yeo, 2007; Gueguen et al., 2013). La norma ISAD-G ja va contemplar el nivell de descripció corresponent a la Unitat Documental Simple (en versió anglesa "*item*"), per tal de recollir la descripció d'aquells elements que no s'agrupaven en grups documentals (Unitats Documentals Compostes, en versió anglesa "*file*"). El fet d'haver triat el concepte d'Unitat Documental Simple, o *item*, enlloc del terme document o *record*, reflecteix ja la sensibilitat envers l'evolució del terme document en el moment de redactar la norma ISAD-G (Niu, 2013; Shepherd & Smith, 2000).

L'entorn anglosaxó ha utilitzat el terme "*Record*" per a descriure els documents amb caràcter d'evidència. El concepte "*Record*" ha estat sempre un concepte genèric, que ha permès encabir tant els documents simples com les agrupacions documentals (expedients i equivalents), i que s'ha adaptat amb molts menys problemes que el concepte "Document" a la descripció de formes documentals diverses (Yeo, 2007, 2011). Les normes originades en l'entorn anglosaxó incorporen aquest concepte amb algunes particularitats. Així, l'*Australian Recordkeeping Metadata Schema* identifica el "*record object*", que diferencia del "*Record aggregation*". La ISO 23081 utilitza també el concepte "*record*", igual que MoReq2010. I el mateix model aplica la norma *Records in Context* (RiC), distingint entre "*Record set*", "*Record*" i "*Record part*".

En canvi, les normes que beuen de la tradició arxivística europea utilitzen el concepte de "document", o sinònims del concepte de document-objecte. Així ho fa l'esquema de metadades de l'e-EMGDE, que utilitza el concepte de "*documento simple*", recollint el model de la ISAD(G), el qual parla de "*item*", i en la traducció castellana de "*unidad documental simple*". També segueix el model ISAD(G) la NEDA, que en aquest cas utilitza directament el concepte de "*unidad documental*".

Una aportació molt interessant de la NEDA, i també de RiC, és que introdueixen un element per sota del concepte de "*unidad documental*", que és el de "*componente documental*" o "*record part*". Recullen amb això l'herència del model PREMIS, que distingeix entre "entitat intel·lectual", "representació" i "component" (cadena de bits o fitxer). RiC contempla a més el concepte de "*instantiation*", equivalent a la "representació" del model PREMIS.

Tornant al discurs inicial, hem destacat que la unitat d'informació es converteix en document només quan adquireix corporeïtat, una dimensió física que el fa gestionable. Hem d'entendre per tant el document com el vehicle de transport de la unitat d'informació, un vehicle que pot ser permanent, temporal o dinàmic (Duranti & MacNeil, 1996). El document, d'acord amb aquesta conceptualització, serveix per transportar la unitat d'informació, no per mantenir-la enregistrada.

La idea de document dinàmic ja apuntada per Luciana Duranti, i la temporalitat del vincle entre la dimensió física i la dimensió conceptual de la unitat d'informació, entronca

directament amb l'evolució en l'ús de la informació digital, i afegeix una nova dimensió al concepte de document: la de document com interacció (Foscarini et al., 2022). El fet d'entendre el document com una interacció d'informació entre agents a través d'uns mitjans tecnològics ha portat a identificar el mitjà o canal d'interacció amb una unitat d'informació, i ha multiplicat les denominacions de les unitats d'informació. Així doncs s'ha parlat de correu electrònic, missatge instantani, contingut Web, etc. com a formes de donar corporeïtat a una unitat d'informació alternatives al document tradicional (Zhang, 2015). Aquesta interacció es fa de forma estructurada (formularis), però també de forma no estructurada (xats, videoconferències gravades, clips de so, etc.), generant noves formes documentals que tindran el mateix valor evidencial que les antigues.

Per l'altra part, les tradicionals dades estructurades en bases de dades relacionals han evolucionat trencant progressivament la dependència dels gestors de bases de dades i adquirint autonomia com a dades estructurades amb una semàntica pròpia, sobretot gràcies als corrents de dades obertes i enllaçades (Borglund & Engvall, 2014). El nou escenari de gestió són per tant dades relacionades de forma dinàmica. Tota expressió tecnològica d'una unitat d'informació, per tant, es retroba actualment en un punt comú en forma d'informació estructurada, tractable en forma de dades independents (i contextualitzades), i integrable en models de dades de caràcter neutre en relació amb la significació individual de cada unitat d'informació, esdevenint la seva consideració com a document una qüestió de perspectiva (Lappin, 2020b).

Es podria afirmar doncs que la dicotomia entre document i dades estructurades s'ha resolt en un solapament de capes, en el qual la capa de dades, conceptualment úniques però que poden estar replicades en diferents fonts, permet generar múltiples representacions, creant unitats d'informació que construeixen significats, els quals posteriorment poden tornar a la capa de dades en forma de noves dades producte de l'acte de creació del significat (Yeo, 2007).⁵⁷

El vincle entre els documents i les dades és doncs cada cop més un vincle circumstancial, dinàmic i subjecte a temporalitat. D'aquesta manera podem definir les dades com una **abstracció del document**, i el document com una **manifestació de les dades**.⁵⁸ Des d'aquest punt de vista, les dades serien quelcom preexistent en l'estructura informacional de l'organització (dades enteses com a variables d'informació),⁵⁹ i per això podem afirmar que extreure l'essència d'un document (la seva abstracció) equival a extreure'n les dades que conté (Yeo, 2007). Els documents els podem entendre com a manifestacions concretes d'aquestes dades, manifestacions

⁵⁷ Aquesta és una de les propostes que es desenvolupen en les conclusions d'aquesta tesi.

⁵⁸ Definició pròpia exposada al I Congrés SEDIPUALB@, celebrat a Albacete el 27 i 28 de setembre de 2022. Vegeu <https://congreso.sedipualba.es/>.

⁵⁹ El terme variables d'informació és utilitzat per la metodologia APS per a designar les dades que contenen els documents (Puig-Pey & Guiu, 2002, 2008, 2013).

creades amb una finalitat concreta (provocar o deixar constància d'una determinada acció) i que tenen com a resultat omplir de valors aquestes variables.

Però als efectes de captura, les dades no són suficients. La creació d'un document a partir de dades genera noves dades, que hauran de retornar a la capa de dades. Però no només això: també genera contextos, és a dir el coneixement que un conjunt de dades s'ha utilitzat en un moment determinat. I això no es captura en forma de noves dades, sinó de relacions entre les dades (Bak, 2016a).

2.2.2 Concepte i tipologia de la dada

El concepte de dada, malgrat la seva recent popularitat, és tremendament polisèmic. La majoria de models actuals el defineixen no de forma absoluta, sinó a partir de les seves implementacions, interpretant que la dada no existeix fora d'un contenidor que l'estructuri, i com a resultat d'aquesta estructuració. I des de l'arxivística no es pot trobar encara, llevat de bibliografia molt recent, referents que formulin una definició "arxivística" del terme dada, llevat d'alguns precedents puntuals (Padilla, 2017; Yakel, 2000).

Devon Mordell es una de les primeres autores a plantejar-se la significació del terme "dada" com objecte del tractament arxivístic, junt amb entendre els arxius com expressió de "Big data" (Mordell, 2019, 2022). Mordell afirma que la dada no és aliena a la professió arxivística, i que també està subjecta, com qualsevol actiu d'arxiu, a un context d'interpretació i a un requisit d'autenticitat per tal d'esdevenir evidència. Però no supera tanmateix la barrera de definir la dada com quelcom vinculat a una implementació tecnològica, i no com una entitat informacional que pugui ser objecte de tractament arxivístic de forma independent al sistema que la conté (Mordell, 2019).

Aquesta idea de dades únicament considerables en un context tecnològic és recurrent en la bibliografia arxivística recent. Padilla afirma que "*To see collections as data begins with reframing all digital objects as data. Data are defined as ordered information, stored digitally, that are amenable to computation.*" (Padilla, 2017). La possibilitat que el tractament a nivell de dada respongui a un replantejament de la metodologia de tractament arxivístic, i no a una acomodació a una realitat d'inel·lable preeminència de la tecnologia digital, com analitza Daniel Rosenberg, és una idea encara poc desenvolupada en l'àmbit dels arxius ("*Data before the fact*", a Gitelman, 2013).⁶⁰

Això no obsta que les dades que custodien els arxius (o el que ja s'anomena els "arxius de dades") s'hagin considerat des de fa temps un recurs a explotar, assimilant-les als recursos naturals amb potencial econòmic (Parilla et al., 2017; Mordell, 2022). S'entenen

⁶⁰ La pregunta de perquè no es pot generar una descripció i una classificació a nivell de dada sense tenir en compte el suport i format en què es trobi aquesta dada encara no té resposta en la teoria arxivística actual.

en aquest cas les dades arxivades com dades en brut (“*raw data*”), dades no tractades de les quals es pot obtenir valor (monetitzable) mitjançant mineria de dades (Moss et al., 2018). Però nombrosos autors discrepen d’aquest plantejament, argumentant que és impossible desconnectar les dades del seu context d’origen, i per tant d’un determinat context d’interpretació o significació no exempt de possibles biaixos ideològics o culturals (Bowker & Star, 1999; Gitelman, 2013; Mordell, 2019, 2022). Fins i tot les dades contingudes dins els documents es comencen a considerar objecte de la descripció arxivística, juntament amb les seves relacions (Parilla et al., 2017).

Comentem a continuació els tipus de dades que la disciplina informàtica ha identificat, i que la ciència de dades utilitza com a camp preferent d’actuació.

2.2.2.1 Tipus de dades segons la informàtica i la ciència de dades

Com s’ha comentat en els apartats anteriors, una distinció habitual que s’ha fet de les dades és entre aquelles dades que no estan pensades per a ser tractades directament com a unitats individuals, sinó com agregats amb significació comuna (dades no estructurades), i aquelles que sí que ho estan (dades estructurades). Malgrat tota la informació digital és potencialment estructurable, només és realment processable quan ha estat convertida, d’una forma o altra, en informació estructurada.

La informació estructurada que podem trobar en una organització es classifica habitualment segons aquesta tipologia (Ghavami, 2020; Laurent et al., 2019):

- Dades transaccionals
- Dades mestres i referencials
- Metadades

Alguns autors proposen d’altres tipus a considerar, com les dades analítiques (Butler, 2011). En aquest estat de la qüestió ens centrarem només en els dos tipus de dades que considerem significatives per al nostre estudi: les dades transaccionals, i les dades mestres.

A continuació aprofundirem en aquestes tipologies de dades, atesa la rellevància que poden tenir per als objectius d’aquesta tesi, doncs creiem que poden incidir en el tipus de sèrie documental i la relació de les sèries amb els sistemes d’informació.

2.2.2.1.1 Dades transaccionals

Cada procés que es gestiona genera dades d'una activitat o transacció, i també d'una sèrie d'objectes o agents que participen en l'activitat. Es consideren dades transaccionals aquelles que descriuen estats operacionals d'una organització (Butler, 2011). Són les dades que manipulen els sistemes d'informació de gestió, també dits sistemes operacionals o, en el cas d'aquesta tesi, transaccionals.

Les dades transaccionals, de la mateixa manera que les dades mestres, necessiten ser consolidades, amb l'objectiu d'eliminar duplicitats i garantir-ne la qualitat. Els entorns que consoliden aquestes dades s'anomenen MDM operacionals, i funcionen de forma similar als MDM que gestionen les dades mestres (Butler, 2011).⁶¹

La bibliografia presta en general menys atenció a les dades transaccionals que a les dades mestres, atès que es consideren el producte "natural" de l'activitat de l'organització, les anomenades "dades en brut" (Ghavami, 2020). Però des del punt de vista arxivístic les dades transaccionals són especialment rellevants, doncs esdevenen un dels potencials punts de contacte entre el concepte de document i el d'informació estructurada (Thomson, 2016).

2.2.2.1.2 Dades mestres

Prestarem més atenció a aquest tipus de dades, atès que són els de més difícil encaixar en la doctrina arxivística i les que, com analitzarem en l'apartat 2.3.3 d'aquest capítol, poden posar en qüestió el criteri de classificació funcional. La bibliografia defineix com a dades mestres aquelles que l'organització està d'acord en considerar com a tals, i que són compartides per tota l'organització (Iqbal et al., 2019). Els típics exemples són client, producte, empleat, etc.

Laurent et al., citant a Gartner, defineixen dades mestres com "el conjunt uniforme i consistent d'identificadors i atributs estesos que descriu les entitats principals de l'empresa, incloent clients, "prospects", ciutadans, proveïdors, llocs, jerarquies i estats de comptes" (Laurent et al., 2019).⁶² En tot cas, les dades mestres es consideren el nucli crític d'informació d'una organització (Laurent et al., 2019).

La gènesi del treball amb dades mestres no és substancial sinó utilitarista, i prové de la necessitat de sincronitzar les dades quan aquestes dades són utilitzades per diferents sistemes d'informació dins una mateixa organització (Iqbal et al., 2019). La manca de sincronia d'aquestes dades pot afectar la consistència dels informes, el compliment de les regulacions o qualsevol projecte que requereixi creuar dades de fonts diverses. Per

⁶¹ Per al detall sobre els sistemes MDM vegeu l'apartat 2.2.3.2 d'aquest capítol.

⁶² On cita a <https://www.gartner.com/it-glossary/master-data-management-mdm/>.

aquest motiu, la identificació de dades mestres va lligada a la posada en marxa de sistemes de gestió d'aquestes dades mestres (MDM) (Kusuma et al., 2019).

Però d'aquest propòsit inicial utilitarista s'acaba derivant un segon propòsit de caràcter holístic: un MDM, un cop implementat, acaba proveint la única versió fiable (de confiança) de la informació d'una organització (Laurent et al., 2019). És el que s'anomena la dada única, o *golden record*.

En molts casos el concepte de dades mestres es defineix únicament per extensió, relacionant una sèrie d'entitats (Laurent et al., 2019; Ghavami 2020). En la majoria d'empreses inclouen quatre categories de dades (Laurent et al., 2019; Ghavami 2020):

- Persones
- Coses
- Llocs
- Conceptes

Per sota d'aquestes categories, sovint s'identifica un segon nivell d'agrupació, conegut com *subject areas*, i també com *dominis* o *tipus d'entitat* (Laurent et al., 2019; Ghavami 2020). La granularitat dels dominis ve determinada principalment pel grau de diferència entre els atributs de les entitats dins cada domini (Ghavami 2020).

Les dades mestres es descriuen (o identifiquen) pels processos de negoci i per la forma com interactuen amb altres dades (Ghavami 2020). I això és precisament el que les posa en relació amb les dades transaccionals. Les relacions es descriuen de la mateixa manera com es descriuen els processos de negoci: un subjecte realitza una acció sobre un objecte (un client compra un producte). Per tant, la relació entre dades mestres i dades transaccionals pot ser vista, segons Ghavami, com una relació entre substantiu i verb. Les dades transaccionals capturen els verbs (és a dir les accions), mentre que les dades mestres capturen els substantius. El contacte amb els plantejaments ontològics és en aquest cas evident (Klischewski, 2006; Quílez, 2011).

Segons Laurent et al. hi ha una certa confusió metodològica en la forma com s'acaba identificant què és una dada mestra i què no és, i no existeix un criteri homogeni (Laurent et al., 2019; Iqbal et al., 2019). Ghavami introdueix alguns elements més subjectius per a la consideració d'unes dades com a dades mestres, com són la major o menor volatilitat (a més volàtils menys mestres), el número de registres (a més registres més mestres) o el valor per a l'organització (a més valor més mestres), lligant el valor a la durada del cicle de vida de la dada (Ghavami 2020). Cal doncs fer una distinció entre la identificació de les dades mestres des d'una perspectiva absoluta, i la identificació de les dades mestres des d'una perspectiva purament funcional. La dada mestra ho és sempre en un determinat context, o només ho és si ha de ser gestionada com a dada mestra? Laurent et al. aprofundeixen en aquesta dualitat, afirmant que malgrat la consideració d'una dada com a dada mestra, només ha de ser tractada (és a dir

gestionada) com a tal quan compleix una sèrie de característiques, amb la qual cosa indirectament restringeix el concepte de dada mestra (Laurent et al., 2019). Les característiques que segons Laurent et al. ha de complir la dada mestra són les següents:

- **Comportament.** Les dades mestres poden ser descrites en funció de com interactuen amb altres dades. En tot procés, les dades mestres acostumen a aparèixer relacionades amb dades transaccionals. I normalment les dades mestres apareixen en més d'un procés, mentre les dades transaccionals són exclusives d'un procés o transacció. Dins un mateix procés, les dades mestres es relacionen amb les dades transaccionals de forma discursiva, en relacions on la dada mestra té el rol de substantiu o nucli, i la dada transaccional exerceix el rol de verb. Per exemple, en la relació un client - compra - un producte, les dades client i producte són dades mestres, mentre que les dades que genera la compra són dades transaccionals.
- **Cicle de vida.** Les dades mestres sovint tenen un cicle de vida propi, un cicle de vida que consta habitualment de quatre fases: creació, lectura-accés, actualització i esborrat (CRUD). Aquestes quatre fases prenen formes i sentits diferents en funció de l'entitat que s'està gestionant, però són molt constants per a totes les entitats mestres.⁶³ Sovint aquestes fases, les accions de creació, accés, actualització i esborrat, són el resultat de diferents processos que, a la seva vegada, generen dades transaccionals. Per exemple, un client es dona d'alta per primera vegada, modifica el telèfon, deixa de ser client, etc.
- **Cardinalitat.** Si entenem la cardinalitat com el nombre d'elements en un conjunt o entitat, com més petit és el nombre d'ocurrències d'una dada, menors són les seves possibilitats d'esdevenir dada mestra. Però determina això només als efectes de disposar d'un sistema de gestió de dades mestres per a la seva gestió, no pas pel que fa a la seva consideració absoluta com a dada mestra. És a dir, les dades mestres d'una organització no seran tractades com a tals (amb un sistema de gestió dedicat) si no arriben a un número Mín. ocurrències que en justifiqui la inversió.
- **Temps de vida (*lifetime*).** Quan parlem del temps de vida, cal distingir entre dos conceptes: el temps de vigència d'una dada (el seu cicle CRUD), i el temps que triga a ser modificada en algun dels seus atributs. Podem anomenar al primer temps de vigència, i al segon volatilitat. No tots els autors fan aquesta distinció,

⁶³ El model de cicle de vida formulat per la ciència de dades afecta una única instància de les dades, i no es considera assimilable al cicle de vida d'una sèrie documental, on els documents i agrupacions passen per diferents etapes on es repeteix l'equivalent al cicle CRUD, motiu pel qual Greg Bak posava en dubte que es pugui definir un cicle de vida unidireccional per a les dades (Bak, 2010).

però considerem que cal aclarir aquest aspecte, atès que serà determinant per concloure l'equivalència entre conjunt de dades i sèrie documental.

- **Temps de vigència.** Les dades mestres tendeixen a tenir un termini de vigència més llarg que les dades transaccionals, i per aquest motiu una mateixa dada pot ser considerada en un context dada mestra, i en un altre context dada transaccional. De fet, les dades mestres tendeixen a ser menys volàtils que les dades transaccionals. Per exemple, un contracte pot tenir una durada plurianual, i per tant una vigència prolongada en el temps, o bé una durada de pocs dies, i per tant una vigència que s'extingeix en molt poc temps. En el primer cas Laurent et al. consideren que hauran de ser tractat com a dada mestra, mentre que en el segon haurà de ser tractat com una dada transaccional.⁶⁴
- **Volatilitat.** Malgrat que les dades mestres tenen menys volatilitat global que les dades transaccionals (són més persistents en el temps i tenen un cicle de vida més llarg), a nivell de detall acostumen a tenir més variabilitat que les dades no mestres (o les que no són tractades com a dades mestres). Aquesta variabilitat es manifesta en la quantitat de vegades que els seus atributs són modificats. Per tant, una entitat els atributs de la qual siguin modificats nombroses vegades al llarg del seu cicle de vigència serà candidata a ser tractada com a dada mestra, mentre aquella entitat amb molt poc moviment en els valors dels seus atributs no ho serà (malgrat des d'un punt de vista conceptual correspongui a una dada mestra).⁶⁵
- **Complexitat.** Les dades mestres (i les que seran gestionades com a dades mestres) tenen un nivell de complexitat elevat, entenent la complexitat com un nombre considerable d'atributs i de relacions que les afecten. Una dada amb pocs atributs serà probablement transaccional.
- **Valor.** Com més valor tingui una dada per a l'organització, major probabilitat de ser considerada (i tractada) com a dada mestra.
- **Reutilització.** El principal motor per a la identificació i gestió de dades mestres és la reutilització. Es defineixen com a dades mestres aquelles que tenen un major potencial de reutilització, i es materialitza aquest potencial a través de la seva gestió mitjançant sistemes dedicats. Si no es fa això, aquest potencial es

⁶⁴ Aplicant un raonament propi de la gestió documental, es podria afirmar que les dades mestres s'haurien de conservar més temps per dos motius. En primer lloc, perquè, com a dades recapitulatives, condensen la memòria de quelcom que ha existit. I en segon lloc, perquè mantenen nombroses relacions narratives amb coses que estan passant en diferents moments, de les quals n'hereten i acumulen terminis.

⁶⁵ Es podria inferir que les dades transaccionals no tenen tanta volatilitat en els valors dels seus atributs perquè són dades discursives, lligades al moment concret en què es desenvolupa una acció, mentre que les dades mestres són dades registrals, lligades a l'evolució d'una cosa.

materialitza de forma natural, mitjançant l'aparició a nivell de procés o d'agent de taules de dades mestres per a ús local, amb els conseqüents problemes de sincronització, actualització i qualitat.

La identificació de les dades mestres és un procés purament tecnològic, però no exempt de vincles amb la gestió documental. Les dades mestres poden adoptar una estructura jeràrquica, i emmagatzemar relacions entre altres dades (Iqbal et al., 2019). Aquestes jerarquies poden ser considerades com un súper domini dins un sistema de gestió de dades mestres (MDM), arran del seu paper crític per poder entendre, i a vegades descobrir, relacions entre dades mestres (Butler, 2011; Kusuma et al., 2019).

Moltes organitzacions no tenen només un set de dades mestres, sinó diversos (Butler, 2011). Això evidencia que el concepte de dades mestres no s'utilitza només per referir-se a dades, sinó també a taules, conjunts de taules o sistemes d'informació. És a dir, se solapen en l'ús el concepte d'aplicació, de conjunt interrelacionat de dades, i de dada elemental.

2.2.2.1.3 Dades referencials

Les dades referencials són un tipus de dades mestres que defineix els conjunts de valors permesos que poden ser utilitzats per altres camps de dades (Butler, 2011). Els típics exemples són països, sexes, classificacions, etc. Es defineixen les dades referencials com "dades que resideixen en taules de codis o *lookup tables*". Normalment són taules estàtiques que emmagatzemen valors com codis de ciutats, d'estats, de districtes, de productes, de països i de classificacions de la indústria. Les dades referencials poden ser definides internament o especificades externament per institucions d'estandardització (ISO, ANSI, etc.), i poden tenir una estructura taxonòmica (Laurent et al., 2019).

Malgrat ser considerades una variant de les dades mestres, les dades referencials se'n diferencien per alguns trets (Sarkar, 2015). El primer és que, mentre que les dades mestres sempre són dades generades dins la pròpia organització, les dades referencials poden ser (i sovint ho són) obtingudes de fonts externes a l'organització. I el segon és que habitualment són més estables i no es modifiquen tant sovint com les dades mestres.

2.2.2.1.4 Diferències entre els tipus de dades

Completem aquest apartat amb una comparativa dels tres tipus de dades que considerem més rellevants per a la posterior equivalència amb el concepte de sèrie documental:

Taula 2-1: Comparativa dels tres tipus de dades

	Transac.	Mestres	Referenc.
L'entitat de dades està emmagatzemada en múltiples sistemes?		X	X
L'entitat de dades fa referència a d'altres entitats de dades?	X	X	
L'entitat de dades categoritza altres entitats de dades?			X
L'entitat pot tenir significat sense fer referència a altres entitats?		X	X
L'entitat pot tenir múltiples atributs associats (a més del codi i el nom/descripció)?	X	X	
L'entitat té un cicle de vida (CRUD)?	X	X	
L'entitat canvia freqüentment en termes de nous registres o canvis en els valors d'atributs? Les dades són volàtils (es creen, modifiquen i esborren molt sovint)?	X	X	
Les dades s'escalen en funció del volum de dades transaccionals?	X	X	
Els nous valors només són creats per canvis en els requeriments de l'organització, i no per canvis en les operacions?			X
L'entitat de dades té un impacte fonamental sobre els processos, els agents i els sistemes?	X	X	

Podem sintetitzar les relacions entre els tres tipus de dades amb la següent afirmació: Els processos de negoci produeixen dades transaccionals. Aquestes dades transaccionals fan referència a objectes o subjectes, que es representen amb dades mestres. I les dades mestres normalitzen les característiques dels objectes o subjectes mitjançant dades referencials.

2.2.2.2 Qualitat i valor de les dades: el valor evidencial de les dades

Sovint el valor evidencial ha estat vist com quelcom exclusiu dels documents, i les dades estructurades s'han vist desafectades d'aquest valor, per adquirir un valor més immediat

i lligat a l'exploració i la reutilització (Khatri & Brown, 2010; Shepherd, 2006). Ha estat més recentment, amb la responsabilitat adquirida per les institucions públiques en la publicació de dades obertes (Borglund & Engvall, 2014; Parycek et al., 2014; Attard et al., 2015; Hellberg & Hedström, 2015), i amb la possibilitat d'autenticar dades amb tecnologia *Blockchain* (Lemieux, 2016; Findlay, 2017; Hofman et al., 2019; Lemieux et al., 2020; Franks, 2020), que les dades s'han començat a veure com un actiu que podia adquirir valor evidencial ple.

La tradició arxivística ha identificat el concepte d'arxiu amb el concepte d'evidència, és a dir d'informació certa. El referent de l'arxivística britànica, Sir Hillary Jenkinson, ho expressava així l'any 1947:

“His Creed, the Sanctity of Evidence; his Task, the Conservation of every scrap of Evidence attaching to the Documents committed to his charge; his Aim, to provide, without prejudice or afterthought, for all who wish to know the Means of Knowledge [...] the good Archivist is perhaps the most selfless devotee of Truth the modern world produces [...]” (Jenkinson, 1980)

L'autenticitat, el valor evidencial, esdevé per tant la raó de ser del concepte d'arxiu, i en conseqüència de la professió arxivística (Magiç, 2008).

Els requisits d'autenticitat requerits als documents electrònics per gaudir d'aquest valor evidencial ja foren formulats a finals dels anys 90 en els projectes de la Universitat de Pittsburg (Bearman, 1994, 2006) i en la primera fase del projecte InterPARES de la Universitat de British Columbia (Duranti & MacNeil, 1996; Duranti et al., 2002). I ja en origen es feia una distinció entre *“trustworthy records as contrasted to decontextualized information or transient data”* (Duranti & MacNeil, 1996).

En l'àmbit de la ciència de dades apareixen de forma menys freqüent els termes propis de la validesa d'un document d'arxiu, i especialment el concepte d'autenticitat (o veracitat de l'origen). Sí que apareixen els termes integritat i fiabilitat, però englobats habitualment dins un conjunt més ampli que s'agrupa sota el terme “qualitat” de la dada (Johnson & Dasu, 2003; Ladley, 2016).

Malgrat el terme no sigui freqüent, el concepte autenticitat no és tampoc absent de la ciència de dades. Ans al contrari, és present a través del concepte de llinatge de la dada (Reina, 2023). Una característica de les dades és que una mateixa dada, un cop dins l'organització i determinada la seva autenticitat, es pot desvincular del document/sistema inicial en què fou capturada (mitjà de transport) i pot passar a altres documents/sistemes sense deixar de ser la mateixa dada ni deixar de tenir el mateix valor (sempre que no es trenqui internament la cadena de custòdia que en garanteix l'autenticitat). El fet que s'incorpori a un sistema de gestió de dades mestres implica conceptualment que es confia en la seva veracitat, i per tant es considera una dada autèntica. L'autenticitat per tant no ve determinada només pel context de la unitat d'informació on ha aparegut per primer cop la dada, sinó també pel control de la cadena

de custòdia de la dada, és a dir per la seva integritat (Findlay, 2017; Järvihaavisto & Öhman, 2022).

Els punts en comú amb la validesa dels documents d'arxiu són doncs evidents, però amb una diferència important: mentre que el document d'arxiu s'entén com una estructura immutable, on l'autenticitat és inherent a la pròpia estructura i no l'hereten per defecte les rèpliques d'aquell document (llevat que es faci un procés d'autenticació del document replicat), l'autenticació en un repositori de dades depèn d'elements contextuals, com la seguretat perimetral o les pistes d'auditoria.

La principal diferència no es troba doncs en la forma com s'entén l'autenticitat de les dades i els documents, sinó en què s'espera d'elles quan es consumeixen. De fet, Terry Cook ja apuntava fa una dècada la *"continuing memory-evidence tension"* en el paper dels arxius, i explicava l'evolució del paper dels arxius de mers custodis de la veritat cap a un paper d'explicar una realitat sempre dinàmica (Cook, 2013):

"The modern conception of memory, therefore, is not something static, not something in the past, not a synonym for "history," and certainly not fixed categories or compartments to aid in memorization or factual recall. Rather, modern memory is an organic dimension of living, an animating series of tools that humans use to make some sense of a rapidly accelerating world of the present day."

També ho apuntava Geoffrey Yeo (Yeo, 2009):

"Sensitivity to communitarian perspectives [...] may oblige archivists to revisit traditional perceptions and extend their understanding of records to encompass new forms of evidence and more fluid manifestations of human memory."

L'analítica de dades basa bona part de la seva confiança en la dada en el coneixement i traçabilitat del llinatge de la dada (*data lineage*), donant per vàlida l'autenticitat verificada en el sistema origen, al qual se li pressuposa veracitat sense tenir sovint control sobre els mecanismes d'autenticació d'aquest sistema (Van Dijck, 2014). De fet, com afirmen Boyd i Jackson, *"the widespread belief that large data sets offer a higher form of intelligence and knowledge that can generate insights that were previously impossible, with the aura of truth, objectivity and accuracy"* (Boyd & Crawford, 2012). Però es demana més: la qualitat de la dada pot implicar a vegades correccions o esmenes de la dada, que fan que no sigui la dada original capturada (Johnson & Dasu, 2003).

La gestió documental tampoc es lliura d'aquesta confiança basada en la presumpció de veracitat, no en la seva demostració. L'àmbit de l'*e-Discovery* ja destaca la irrellevància "de facto" de la distinció arxivística entre *"record"* i *"document"*, atès que en qualsevol procés judicial i als ulls de la Llei tot és potencial evidència (Lueders, 2021). Chris Walker afirma que, des del punt de vista de la governança (holística) de la informació (Walker, 2023):

“[...] no distinction is drawn between records and non-records. From a business execution perspective the difference is irrelevant, from an evidentiary perspective it's minimal since any information you have can be used against you in proceedings.”

L'arxivística ha fet avenços conceptuals importants per garantir l'autenticitat de la informació en entorns distribuïts, on no existeix un control centralitzat, i on els documents no disposen d'una autenticació autònoma com és la signatura electrònica basada en certificat. Ho ha fet sobretot a partir dels treballs impulsats per la Universitat de *British Columbia* i en el marc del projecte *InterPARES Trust* (iTrust)⁶⁶ amb la utilització de *blockchain* com a sistema d'autenticació documental (Lemieux, 2016; Findlay, 2017; Hofman et al., 2019; Lemieux et al., 2020; Franks, 2020; Duranti, 2022).

Un corrent important dins les organitzacions és el que alguns autors defineixen com a democratització de (la gestió de) les dades (Labadie et al. 2020), i que en l'entorn empresarial rep el nom de *data mesh* (Dehghani, 2019; Delgado, 2022). Labadie et al. consideren que aquest corrent beu de la idea dels principis FAIR, proposats per Wilkinson et al. (Wilkinson et al. 2016), y que els propis principis FAIR, originats en l'àmbit de les dades científiques, s'estan estenent cada cop més a l'àmbit empresarial. Cada terme dels continguts en l'acrònim FAIR (*Findability, Accessibility, Interoperability* i *Reuse*) es vincula a un conjunt de principis sobre els requisits que han de complir les dades en un entorn adequadament gestionat. El vincle amb el model *data mesh* prové del fet que FAIR no pressuposa cap estructura organitzativa per a la gestió de les dades, i fa possible delegar en el productor o custodi de les dades (*Data steward*) la responsabilitat que les dades compleixin els principis FAIR.

De fet ja trobem un antecedent del model *data mesh* en la designació de responsabilitats sectorials que el mètode *Capstone* del NARA proposa per a les Agències Federals. Ho personifica en el que anomena *Capstone officials*, que són els càrrecs d'un ens responsables del desenvolupament de polítiques i programes en aquell ens que són designats per vetllar per l'aplicació de la GRS 6.1 (*White Paper on The Capstone Approach and Capstone GRS*, 2015).

Això ens porta a convenir que en la gestió de dades s'està produint una evolució sobre la propietat de les dades determinant per al concepte de sèrie documental, atès que afecta un dels seus components principals: l'agent productor.⁶⁷ Així doncs, s'evoluciona (Khatri & Brown, 2010; Dehghani, 2019):

⁶⁶ Vegeu <http://interparestrust.org/trust>.

⁶⁷ I, en conseqüència, posa en qüestió el tradicional principi arxivístic de procedència.

Taula 2-2: Evolució en la responsabilitat sobre les dades

Del **propietari** de les dades (*data owner*), que es limita a produir i custodiar les dades, però sense cap responsabilitat “orgànica” sobre la qualitat d’aquestes dades, cosa la qual genera sovint dades d’interpretació captiva (no documentades o sense metadades) i de qualitat dubtosa o no verificable.

Al **responsable** de les dades (*data steward*), un productor al qual l’organització ha designat com a “responsable” de la qualitat de les dades, que respon corporativament d’aquesta qualitat. Això permet que la gestió de les dades i de la seva qualitat es faci de forma descentralitzada, compartint la responsabilitat de la seva qualitat entre tots els productors.

El sistema de governança de la informació passa en aquest cas de la preocupació per la qualitat de les dades finals (cosa la qual representa un alt cost per garantir aquesta qualitat, que inclou *data cleansing*, *data deduplication*, etc.) a donar suport als *data stewards* per a que siguin ells qui vetllin per aquesta qualitat, de forma descentralitzada (Ladley, 2016). L’ús de tècniques de visualització de dades, de quadres de control, i d’automatismes que facin verificacions semiautomàtiques d’aquesta qualitat és una forma que contribuir a reforçar el paper dels *data stewards* com a corresponsables de la qualitat de les dades del conjunt de l’organització (Lemieux et al., 2014; Lemieux, 2015).

2.2.2.3 La descripció de les dades: els catàlegs de dades

Dins les organitzacions les dades estan emmagatzemades en diferents sistemes i en formats heterogenis. L’harmonització d’aquestes dades és necessària per poder avançar en el seu ús, per exemple utilitzant-les per aplicar tècniques d’intel·ligència artificial. I per assolir aquesta harmonització, es considera necessari organitzar i classificar aquestes dades mitjançant la inclusió en un catàleg de dades que permeti assignar-hi metadades (Ehrlinger et al. 2021). Per aquest motiu, els anomenats catàlegs de dades s’estan convertint en l’instrument de descripció de referència, que permet emmagatzemar metadades sobre les dades.

El concepte de catàleg de dades va popularitzar-se a partir d’un informe de Gartner (Zaidi et al. 2017), qüestionat posteriorment per altres autors (Bullivant 2018). El catàleg de dades s’ha anat configurant com l’eina que ha de centralitzar en una organització la recollida i gestió de les metadades sobre les dades, com a recurs bàsic per a cerca i explotació (Quimbert et al. 2020). Es planteja com una eina que va més enllà de la integració de dades provinents de diverses fonts, sinó que a més les enriqueix semànticament a través de la metadescripció (Ehrlinger et al. 2021).

L’assignació de metadades a les dades es considera cada cop més una necessitat per a la seva gestió integrada. L’oferta de programari per a la implementació del catàleg de

dades és cada cop més àmplia, i habitualment prové de productors amb experiència en tecnologia de base de dades (*Oracle, Informatica, IBM, Denodo, Alation, etc.*).

Però malgrat la seva operativització, s'evidencia encara una manca important de recerca i referents metodològics per a la construcció i normalització de catàlegs de dades (Ehrlinger et al. 2021, Labadie et al. 2020). De fet, l'anàlisi de Ehrlinger et al. només identifica, des de l'any 2000, un total de 11 publicacions que facin aportacions rellevants, des d'una òptica transversal, per a la definició, composició i metodologia de desenvolupament de catàlegs de dades. Malgrat l'ús del terme catàleg de dades és cada cop més estès, en la majoria de casos es tracta de repertoris temàtics o sectorials de *datasets*, especialment per donar suport a la recerca en determinats àmbits científics, i poques vegades s'utilitza el terme per a descriure un instrument propi d'un sistema de govern de la informació o de la dada (Ehrlinger et al. 2021). De fet, Ehrlinger et al. no identifiquen cap definició per extensió de catàleg de dades, i només poden extreure quatre elements com a denominador comú: la gestió de metadades, el context de negoci, les responsabilitats i rols sobre les dades, i els principis FAIR. En la mateixa línia, Labadie et al. consideraven que el concepte de catàleg de dades és encara massa nou i poc definit, i que manquen estudis empírics de la seva implementació (Labadie et al. 2020). Les definicions donades fins el moment dels catàlegs de dades, com l'original de Zaidi et al. en la qual defineix catàleg de dades com l'instrument que "*maintains an inventory of data assets through the discovery, description, and organization of datasets*" (Zaidi et al. 2017) resulten probablement limitatives en relació amb el funcionament actual de les plataformes de govern de la dada.

Ens centrarem doncs en el primer component identificat per Ehrlinger et al., la gestió de les metadades. La gestió de metadades en un catàleg de dades té en principi molts punts en comú amb la gestió de metadades de qualsevol altre tipus d'actiu digital, així com la forma de captura, manual o automatitzada (Shanmugam i Seshadri 2016). El tret diferencial cal cercar-lo en la informació recollida en les metadades sobre les dades. Alguns autors refereixen a models clàssics (Quimbert et al. 2020), diferenciant entre metadades descriptives, administratives i estructurals (Riley 2017). Però d'altres, com Srinivasan Shanmugam i Gokul Seshadri proposen models més orientats a la gestió específica de dades i *datasets* (Shanmugam i Seshadri 2016):

- Dins les metadades descriptives:
 - Qualitat
 - Fiabilitat
- Dins les metadades administratives:
 - Llinatge
 - Context tècnic
 - Sensitivitat i accessibilitat
- Dins les metadades estructurals:

- Relacions amb sistemes
- Relacions entre dades
- Context de negoci

En tots els casos, la provenença de les dades (l'linatge, relacions amb sistemes, etc.) esdevé un dels elements fonamentals del context de les dades, i base del paper que juguen els catàlegs de dades com a l'instrument que agrega dades de múltiples fonts (Vicknair 2010). Els altres tres elements que segons Ehrlinger et al. Conformen un catàleg de dades, és a dir el context de negoci, les responsabilitats i rols sobre les dades, i els principis FAIR (Ehrlinger et al. 2021), s'acaben traduïnt en elements d'aquest esquema de metadades (Shanmugam i Seshadri 2016).

La catalogació de dades no és lineal, i hereta molts elements funcionals propis del modelat de dades, cosa la qual ha portat a alguns autors a proposar les ontologies com la forma més expressiva per construir un esquema de metadades orientat a descriure dades (Feilmayr i Wöß 2016, Quimbert et al. 2020). En aquesta línia, s'han utilitzat per a la catalogació de dades ontologies com DCPAC (*Data Catalog Provenance, and Access Control*), combinada amb DCAT (*Data Catalog Vocabulary*)⁶⁸ i amb PROV-O (PROV Ontology).⁶⁹

L'assignació de responsabilitats i rols sobre les dades, un element crucial per garantir-ne la fiabilitat (Shanmugam i Seshadri 2016), és alhora un element de l'esquema de metadades i un model de gestió de les dades, que ha anat configurant la figura del responsable sectorial de dades o *data steward* com a rol clau per garantir la qualitat, fiabilitat i actualitat de les dades (Labadie et al. 2020). Aquest element, clau per al compliment dels principis FAIR, i per a la gestió de les dades en un entorn de responsabilitat distribuïda (*data mesh*), es converteix doncs en un dels elements de metadades més importants.

Els propis principis FAIR conformen una guia de quines metadades ha de contenir un catàleg de dades, a partir de les característiques que han de complir les dades i que han de quedar consignades en forma de metadades (Redkina, 2019). Podem trobar una anàlisi detallada d'aquest vincle en les nou funcions que Labadie identifica en els catàlegs de dades (Labadie et al. 2020).

La relativa migradesa bibliogràfica en la descripció dels components d'un catàleg de dades també afecta les propostes sobre la seva construcció i desenvolupament (Ehrlinger et al. 2021). Només tres propostes identifica Ehrlinger, i totes coincideixen en considerar la construcció d'un esquema de metadades com el principal pas en la construcció d'un catàleg de dades (Wang 2014, Shanmugam i Seshadri 2016).

⁶⁸ Vegeu <https://www.w3.org/TR/vocab-dcat>.

⁶⁹ Vegeu <https://www.w3.org/TR/prov-o>.

Shanmugam i Seshadri proposen un procés en vuit passos per a la construcció del catàleg:⁷⁰

1. Definir el context de les variables de dades (relacions de les dades amb sistemes d'informació, amb processos, amb tecnologies, i el llinatge de les dades).
2. Definició dels atributs de les dades (qualitat, sensitivitat, accessibilitat i fiabilitat).
3. Etiquetat (*tagging*) de les dades, per tal de decidir quines metadades apliquen a cada nivell de descripció (columna, entitat o *dataset*).
4. Definició de regles per controlar l'accés a les dades i auditar-lo.
5. Assemblatge de l'esquema final del catàleg de dades en un model de dades concret (relacional, ontològic, etc.).
6. Emplenament del catàleg de dades amb dades.
7. Presentació/exposició del catàleg de dades al públic.
8. Recollida de revisions i valoracions que permetin millorar el catàleg.

Labadie et al. fan dues aportacions metodològiques importants en relació amb la construcció del catàleg de dades (Labadie et al. 2020), diferenciant entre formes d'aproximació. Pel que fa a la construcció de l'esquema de metadades, les dues formes d'aproximació són:

- *Top-down*: primeres es defineix l'estructura, i després s'importen les dades.
- *Bottom-up*: l'esquema es construeix a partir de l'anàlisi de les dades importades.⁷¹

I pel que fa a la seva implementació, Labadie et al. distingeixen entre dues aproximacions:

- *Data supply-driven* (o *input-oriented*): en el qual es fan peser sobretot els requeriments dels usuaris que dipositaran dades en el catàleg.
- *Data demand*: en el qual el focus es posa en els requeriments dels usuaris del catàleg i en quines hauran de ser les sortides del catàleg.

⁷⁰ Els cinc primers punts d'aquest model descriuen la construcció d'un instrument d'ús general, o vàlid per a la governança, mentre que els tres darrers ja impliquen la construcció d'un producte concret d'informació. En un entorn científic, on el catàleg es construeix per facilitar l'explotació de les dades, cal generar un producte d'informació acabat, i per tant prenen sentit els vuit punts de Shanmugam i Seshadri. Però en un entorn empresarial, i especialment quan l'objectiu del catàleg és esdevenir un sistema de governança, els cinc punts inicials ja són suficients per disposar d'un instrument suficientment operatiu.

⁷¹ Aquesta segona aproximació és la que és més fàcil d'automatitzar, i també la que hem adoptat en el cas d'estudi del Consorci d'Educació de Barcelona, en el qual el mapa de dades s'ha construït de forma incremental, a partir de la identificació progressiva de dades en evidències i sistemes d'informació, i de la seva contextualització a partir dels objectes descrits.

En tot cas, l'ús previst del catàleg de dades és el que determinarà la forma que es prioritza. I en l'entorn empresarial, on el que es busca és potenciar la interoperabilitat interna, tant a nivell de sistemes d'informació com de processos, el paper del catàleg esdevé fonamental per fixar consens en relació amb el mapa i significació de les dades, i per determinar les responsabilitats sobre les dades (Kurth, Ruddy, i Rupp 2004).

El punt crític per a l'arxivística i la gestió documental el trobem en la granularitat de la descripció. La gestió de metadades implica, com en qualsevol model de descripció, gestionar dades a diferents nivells de descripció (Chen 2022). Mentre aquests nivells estan clars en els models arxivístics, en les experiències de descripció de dades estructurades els autors utilitzen una nomenclatura molt genèrica, com "recurs d'informació" o "actiu d'informació". I en desgranar-ho en nivells, observem que descriuen:

1. Bases de dades (*datasets* o conjunts de dades)
2. Taules de base de dades (també denominades sovint *datasets*, o "entitats")
3. Columnes o camps de base de dades

Els atributs amb els que habitualment es descriuen aquests elements són els següents (Ghavami 2020):

- Nom o identificació
- Descripció o definició
- Descriptors
- Font de les dades i història dels moviments i transformacions al llarg del seu cicle de vida ("*data lineage*" o llinatge de les dades)
- Registre de qui crea les dades i qui consumeix les dades ("Tracking usage" o traçabilitat), siguin aquests agents humans o aplicacions
- Relacions amb productes, agents i processos de negoci, definint quines dades utilitzen quins agents en cada procés i per a prendre quines decisions
- Informació sobre privacitat, confidencialitat i accés

Existeixen nombrosos models de metadades per a la descripció de dades en *data lakes* (Wieder & Nolte, 2022), vinculats a diferents models en l'estructura de l'emmagatzemament, com GEMMS (Quix et al., 2016), MEDAL, GoldMEDAL i CODAL (Sawadogo & Darmont, 2021), Network-based (Diamantini et al., 2018) GOODS (Halevy et al., 2016) o CONSTANCE (Hai et al., 2016).⁷²

⁷² L'Administració General de l'Estat ha aprovat la *Resolución de 19 de febrero de 2013 (BOE de 4 de marzo)*, de la *Secretaría de Estado de Administraciones Públicas*, por la que se aprueba la *Norma Técnica de Interoperabilidad de Reutilización de recursos de la información*, la qual utilitza

La majoria de catàlegs es visualitzen com eines independents, desconnectades d'altres àmbits de la gestió de la informació (gestió documental, gestió dels sistemes i aplicacions, etc.). Per a gestionar aquesta metainformació cal disposar d'eines de gestió de metadades, també anomenades "*Metadata hub*" (MDH) o "*Metadata registry*".

2.2.3 Els sistemes d'informació

Tradicionalment l'arxivística ha entès el sistema d'informació com una eina per a la creació de documents (Bearman, 1991; Duranti & MacNeil, 1996; *Committee on Electronic Records*, 1997; Bantin, 1998; Duranti et al., 2002). Només en alguns casos el sistema d'informació s'ha considerat com un document en el seu conjunt, i en molts pocs casos s'ha intentat establir un vincle sistemàtic entre els sistemes d'informació i les entitats o agrupacions documentals (Suderman, 2002). Recentment aquest debat s'ha diluït encara més a causa dels nous models de sistemes, les modalitats de servei SaaS i la separació progressiva d'una capa de dades cada cop més autònoma semànticament dels components executables de les aplicacions (Duranti, 2022).

Els casos en què els sistemes d'informació han centrat l'interès de l'arxivística han estat quan podien convertir-se en un instrument útil per a desenvolupar-hi els processos arxivístics de classificació, descripció, conservació, eliminació, etc. Això ha afectat principalment dos tipus de sistemes: els anomenats gestors documentals (EDMS) (Vieira et al., 2012; Sataolaatten, 2014) i els repositoris per a la preservació digital (Wilson, 2017). I també ho han fet puntualment alguns entorns amb característiques pròpies, com el correu electrònic o les pàgines Web (Saffady, 2014).

Però en aquesta tesi considerem que cal entendre el sistema d'informació (i especialment l'aplicació informàtica de negoci) com una representació operativa d'un context funcional, i per tant un element que dona context (trasllada aquest context) als documents i dades que genera. No és un element neutre; és un element al qual donarem nom en funció dels processos als quals dona suport, i en definirem els límits també des d'aquesta perspectiva funcional. Això ja ens permet apuntar un primer paral·lelisme entre els sistemes d'informació i les sèries documentals, que es desenvolupa en la conclusió de l'objectiu 3 (apartat 8.1.3 del capítol 8).

Actualment, però, el principal focus d'atenció en la gestió de la informació organitzacional són les dades i els repositoris on rauen. Quan es parla de dades, sempre es pensa en les bases de dades que les contenen com quelcom també neutre, i no en els sistemes que les generen en un context funcional concret. Una base de dades és un constructe tecnològic. Però aquest constructe es deriva d'una necessitat de consum de dades per part d'una aplicació informàtica, i aquesta aplicació es deriva d'unes

DCAT com a estàndard de descripció de conjunts de dades. Disponible a https://www.boe.es/diario_boe/txt.php?id=BOE-A-2013-2380.

necessitats funcionals generades per un procés de negoci. L'estructura de la base de dades i la semàntica de les dades no es poden desconnectar dels contextos que les generen ni dels que les utilitzen. Hem d'entendre, per tant, que una base de dades no és mai un element neutre, sinó quelcom construït amb una intencionalitat concreta, que sovint es reflecteix en el modelat de les dades que conté (McLeod, 2008; Lappin, 2010).

Aquesta absència del concepte de sistema d'informació es fa palesa en les normes de descripció arxivística. Ni a la norma ISAD-G, ni a les ISO 23081 (i models previs i derivats), ni a la NEDA, no es fa esment dels sistemes d'informació dins la panòpia d'entitats a gestionar per part d'un sistema de gestió documental, i en l'apartat d'agents només es parla específicament de persones. Sí que en fan esment la e-EMGDE (que parla de "*Persona o dispositiu*") i la norma *Records in Context* (RiC), però en ambdós casos ho assimila a l'entitat agent, i no proposa considerar-ho com una entitat a part o assimilada a cap tipus d'agrupació documental.

Les dades, quan no formen part de documents-objecte, només existeixen dins un ecosistema tecnològic. Aquest fet ha determinat la seva vinculació a la disciplina informàtica, com també la lògica amb la qual s'han estructurat per al seu processament i explotació. Aquesta dependència també ha estat determinant en la dicotomia entre el model docucèntric, on l'objecte és el document, i el model datacèntric, on, enlloc de la dada, ha estat el sistema en el seu conjunt el que s'ha equiparat amb el document (Harries, 2000, 2009, 2012). Ha estat a partir d'aquí que s'ha generat la idea de considerar la base de dades una entitat amb significació arxivística, com també qualsevol component que faci funcions de base de dades. I d'aquí també es deriva la tendència actual a considerar el concepte de conjunt de dades (o *dataset*) com l'entitat a gestionar quan es planteja la governança o la conservació de les dades (Shepherd & Smith, 2000; Niu, 2016; Campos, 2023).

Per aquest motiu, i amb la voluntat d'escatir la pertinència d'aquestes equivalències, volem detallar breument els tipus, tecnologies i arquitectures més utilitzats actualment per a l'emmagatzematge i processament de dades. Analitzem doncs a continuació els diferents models i entorns d'emmagatzematge de dades, i cerquem punts de contacte amb el concepte de sèrie o agrupació documental.

2.2.3.1 El model relacional com a referent dels sistemes de gestió

Les bases de dades relacionals són objectes digitals complexos amb una estructura de dades acuradament definida (Ur et al. 2015). Estan basades en els fonaments formals del model relacional proposats als anys 70 per E. F. Codd (Codd 1970, Codd 1971), i impliquen un procés de normalització de la informació, segons una estricta progressió de formes normals (Kent 1983), que assegura la unicitat de la dada i la identificació

unívoca de les ocurrencies de cada entitat en el conjunt de la base de dades a través de la clau primària. El resultat d'aquesta normalització és una estructura de taules altament consistent i adaptada a la captura de fets i accions, però notablement fragmentada i difícilment assimilable de forma directa a entitats significatives del món real (Ur et al. 2015), i especialment a entitats amb significació documental.

És per aquest motiu que la literatura arxivística ha afirmat sovint que una base de dades (entesa com una base de dades relacional) és una estructura merament tecnològica sense possibilitat de significació documental (Yeo, 2007, 2011, 2012).

No obstant, aquest no és motiu per negligir la transcendència del model relacional en la gestió d'informació amb valor arxivístic (Heuscher et al., 2004). El model relacional segueix essent la base de la major part de les aplicacions de gestió en les organitzacions actuals, i una part molt important dels sistemes que es desenvolupen actualment, especialment al sector públic, es basen únicament en models relacionals. Aquest model segueix essent el referent pel que fa a consistència de les dades, i també com a garantia de precisió en el tractament de dades on no és tolerable cap grau d'incertesa o variabilitat.

2.2.3.2 Els sistemes de gestió de dades mestres (MDM)

El problema de disposar d'una mateixa dada, descriptora d'un mateix element, instanciada amb valors diferents en diferents sistemes d'informació es la raó per la qual cal posar en marxa un sistema de gestió de dades mestres (Butler, 2011). Laurent et al. defineixen gestió de dades mestres (*Master Data Management*, o MDM) com la tecnologia, eines i processos requerits per crear i mantenir llistes consistents i acurades de dades mestres (Laurent et al., 2019). Aquesta definició inclou alguns elements a ser tinguts en compte. El primer és que l'MDM no resol només un problema tecnològic, sinó també organitzatiu, doncs certs processos poden requerir la creació o captura de dades mestres, i això pot requerir a la vegada canvis en els processos. I el segon és que l'MDM inclou no només crear, sinó també mantenir les dades mestres, atès que l'esforç de depurar i consolidar unes dades com a dades mestres no es justifica si no és possible després actualitzar-les de forma que es mantingui la seva consistència (Iqbal et al., 2019).

Però la inversió per posar en marxa un sistema MDM no es justifica de forma immediata ni en tots els casos, per la qual cosa alguns autors recomanen començar desplegant només un nombre limitat de dades mestres, i fer créixer progressivament el model a mesura que se'n vagi justificant la inversió (Kusuma et al., 2019).

Hi ha tres opcions arquitecturals per implantar en una organització un sistema de gestió de dades mestres (Laurent et al., 2019):

Taula 2-3: Models d'arquitectura per a la implantació d'un sistema de gestió de dades mestres

Sistema	Funcionament	Punts forts	Punts febles
Còpia única (exemplar únic)	Només hi ha un exemplar de les dades mestres. Totes les altes i modificacions es fan directament sobre aquest exemplar. Totes les aplicacions es reprogramen per utilitzar aquesta font de dades enlloc de la pròpia.	Es garanteix la consistència gairebé absoluta de les dades mestres.	La implantació és molt costosa en temps i cost econòmic, ja que cal modificar totes les aplicacions, i quan es treballa amb aplicacions de mercat (adquirides) o al núvol a vegades és impossible.
Múltiples còpies amb manteniment centralitzat	Només hi ha un exemplar de les dades mestres. Totes les altes i modificacions es fan directament sobre aquest exemplar. Cada aplicació té una còpia íntegra o parcial de la taula mestra, que és disseminada periòdicament des de la taula mestra. Però les aplicacions no poden modificar les dades mestres, només llegir-les i vincular-hi noves dades.	Es garanteix la consistència gairebé absoluta de les dades mestres. I no cal fer tantes modificacions a les aplicacions de gestió com en el model anterior.	La implantació requereix igualment la modificació de les aplicacions existents, i bloquejar-ne algunes funcionalitats per evitar la modificació de les dades mestres.
Fusió continuada (<i>continuous merge</i>)	Hi ha un exemplar de la dada mestra, que es dissemina creant una còpia íntegra o parcial a cada aplicació. Qualsevol aplicació pot realitzar canvis sobre la taula mestra, que automàticament es disseminen a la resta de còpies de la taula mestra a cadascuna de les aplicacions.	Es redueix la quantitat de canvis que cal fer en les aplicacions de gestió, ja que només cal resoldre la missatgeria entre aplicacions (capa de serveis).	Mantenir la consistència de les dades és més difícil per tres motius: <ol style="list-style-type: none"> 1. Cal preveure mecanismes de resolució de conflictes per modificació concurrent d'una mateixa dada en la taula mestra, i estendre les validacions de qualitat de les dades mestres a cadascuna de les aplicacions que les poden modificar. Aquests mecanismes poden implicar l'ús de regles de governança de dades (prioritat) o la intervenció manual d'un <i>data steward</i> per a resoldre certs conflictes. 2. L'addició concurrent pot crear duplicats. Cal mantenir un sistema de <i>matching</i> per a detectar i eliminar aquests duplicats. 3. Cal preveure canvis que afectin la consistència de les dades (p.e. canvis en les unitats de mesura) i que poden ser vistos al moment pels usuaris, creant confusió sobre la fiabilitat de la dada.

La posada en marxa d'un MDM és un procés complex amb un temps de desenvolupament llarg (Kusuma et al., 2019). La clau per a una correcta implementació d'un MDM és, en primer lloc, una implementació progressiva i incremental, que permeti controlar-ne els aspectes de detall i oferir beneficis evidents en el menor temps possible. I en segon lloc, comptar en tot moment amb els usuaris coneixedors de cada domini de negoci, sense la participació dels quals és impossible disposar del coneixement suficient per interpretar i validar correctament la qualitat de les dades.

2.2.3.3 Els models orientats a l'exploració: el modelat dimensional i el *business Intelligence*

El concepte d'intel·ligència de negoci (o *business intelligence*) fou formulat per primer cop l'any 1958 per Hans Peter Luhn, investigador de IBM, i posteriorment, el 1989, popularitzat per un analista de Gartner, Howard Dresner (Julbe, 2023). L'exploració de la informació requereix transformacions intenses de l'estructura de les dades per respondre a les preguntes de l'organització. Aquestes transformacions, quan els volums de dades són limitats, es poden fer eficaçment amb el model relacional. Però el processament de grans volums de dades penalitza el rendiment del sistema, i en determinats contextos fa inviable l'exploració (Khine & Wang, 2018). Per aquest motiu en l'àmbit de la intel·ligència de negoci (*business intelligence*) van aparèixer els *data warehouse*.

En un *data warehouse*, l'estructura de les dades segueix regles diferents a les d'un model relacional. Mentre en un model relacional s'evita la repetició o duplicació de cap dada, en un *data warehouse* les dades es desnormalitzen (en relació amb les formes normals de Codd) per construir taules tridimensionals (taules de taules), els anomenats cubs d'informació (*OnLine Analytical Processing*, o OLAP) o més endavant *datamarts* (Julbe, 2023).

La desnormalització es pot fer amb diferents tècniques, i consisteix habitualment en passar d'un modelat relacional a un modelat dimensional (Ur et al., 2015). L'esquema en estrella agrupa les dades en dimensions, corresponents a objectes o entitats pròpies de dades mestres (coses del món real), que es relacionen amb una entitat pròpia de sistemes transaccional (activitat o transacció), que es recull en el que es denomina taula factual ("*fact table*") (Butler, 2011). El conjunt d'una taula factual i les dimensions involucrades en el procés descrit per la taula factual és el que s'anomena una estrella (Ur et al., 2015). En aquesta tesi ens hem inspirat en l'esquema en estrella per configurar el model ontològic de classificació arxivística de les dades (objectes "cosa" i objectes "acció").

2.2.3.4 Els models distribuïts: el model llac de dades i el *Big Data*

L'aparició cap a finals dels 90 de les bases de dades NoSQL (*Not Only SQL*), i la seva facilitat per a desplegar-se de forma distribuïda, amb un consum racional de recursos computacionals (especialment amb tecnologies com *Hadoop*, *Spark* i *MapReduce*), va obrir la porta a un nou tipus d'arquitectures d'emmagatzemament de dades (Julbe, 2023). Aquestes arquitectures no podien complir amb els estrictes models de consistència de les bases de dades relacionals i dels *data warehouse*, com ACID (*Atomicity, Consistency, Isolation and Durability*), formulat per Haerder i Reuter (Haerder & Reuter, 1983), i utilitzaven regles més flexibles, com el model BASE (*Basically Available, Soft state, Eventually consistent*) (Khine & Wang, 2018).

En una arquitectura distribuïda, les dades s'emmagatzemen en múltiples repositoris, la majoria de vegades sense una significació documental concreta, i es converteixen en agregats significatius quan són recuperades i utilitzades per al negoci (Julbe, 2023). L'impacte d'aquestes construccions sobre la gestió documental i l'arxivística és evident.

El concepte de llac de dades (*data lake*) fou proposat per James Dixon, CEO de Pentaho, l'any 2010 (Dixon 2010). La voluntat d'alliberar la gestió de les dades dels rígids esquemes del model relacional ja s'havia fet palesa als anys 80 amb l'aparició de les bases de dades orientades a objectes, i ha continuat evolucionant fins a la revolució del *Big Data* (Chen 2022). La principal diferència amb els models anteriors és la capacitat d'admetre, juntament amb la informació estructurada, altres fonts d'informació semiestructurada o directament no estructurada (Khine & Wang, 2018).⁷³

Això és possible per la flexibilització de la forma i el moment en què es processa la informació. Mentre el model propi del *data warehouse* es basa en estructurar la informació d'acord amb un esquema rígid en el moment en què ingressa en el sistema (*schema on write*), en un *data lake* la informació s'estructura directament en el moment en què es consumeix (*schema on read*) (Chen 2022; Wieder i Nolte 2022). Així doncs, a diferència del procés de tractament inicial propi dels *data warehouse*, basat en l'execució de processos de transformació de tipus ETL (*extract, transform and load*), en un *data lake* l'ordre de processament canvia. Mentre en el model relacional o en un *data warehouse* les dades només es gestionen quan ja han estat estructurades segons un determinat esquema (*conformed data*), en un *data lake* les dades s'emmagatzemen en el seu format original en brut (*raw data*), i el pre-processament d'aquests dades només es porta a terme quan és requerit per alguna aplicació o per una consulta d'usuari. Per

⁷³ Des d'aquest punt de vista, un *data lake* pràcticament no es diferencia del concepte d'arxiu digital, atès que pot emmagatzemar i governar pràcticament qualsevol tipus d'informació digital. No només això, sinó que a més es concep com un repositori d'informació original, no necessàriament tractada en el seu procés d'ingesta. Els conjunts de dades (*dataset*) que ingesta un *data lake* provenen directament dels sistemes o aplicacions que els generen, són dades en brut (*raw data*).

aquest motiu en l'entorn dels *data lake* es considera que el procés ETL es transforma en un procés ELT (*extract, load and transform*) (Khine & Wang, 2018; Waddington et al., 2021).

Un *data lake* utilitza una estructura plana, sense cap mena de jerarquia entre els continguts, on cada element es descriu amb un identificador únic i amb un conjunt de metadades. L'únic ordre que mantenen els elements és l'ordre d'arribada (la data d'ingesta), per tal de poder disposar d'una lectura diacrònica correcta de les dades. Això fa que l'actualització de dades en un *data lake* sigui més àgil que en un *data warehouse* (on l'actualització depèn d'una nova execució de les ETL), i que en conseqüència un *data lake* pugui disposar més fàcilment de dades actualitzades, i en major volum que un *data warehouse* (Chen 2022).

El procés d'explotació de dades en un *data lake* s'anomena exploració (Makhlouf Shabou et al., 2020). A diferència del procés d'explotació en un *data warehouse* (*data mining*), i donat el volum i heterogeneïtat de les dades que conté, l'exploració en un *data lake* no pren una forma tant lineal. Així doncs, la consulta i anàlisi de les dades d'un *data lake* pot orientar-se a la resposta a una pregunta en concret (obtenir una mesura o identificar un patró), però molt sovint té per objectiu localitzar dins el *data lake* aquells *datasets* que contenen informació rellevant per respondre a les necessitats d'informació d'un usuari, i que seran posteriorment processats de forma individual per respondre a preguntes concretes. Aquest nou paradigma d'accés a la informació és el que s'anomena *query-driven data discovery* (Diamantini et al. 2018; 2021). De totes maneres, malgrat la teoria d'un *data lake* és disposar de les dades abans de fer-hi cap processament, a la pràctica és inevitable algun tipus de tractament durant la ingesta per a fer possible conèixer i poder explotar les dades ingestades (Chen 2022).

Els principals reptes actuals per als *data lake* són els que afecten a l'arquitectura i al govern de les dades, especialment aquells relacionats amb la catalogació i metadescripció dels *datasets* (Chen 2022). Això fa que un element cabdal en un entorn amb esquemes flexibles sigui el control del cicle de vida de la dada, en el qual s'inclouen els processos de *data ingestion*, *data extraction*, *data cleaning*, *dataset recovery*, *metadata management*, *data integration* i *dataset versioning* (Chen 2022).

El model *data lake* s'estructura habitualment en tres capes funcionals (Hai et al., 2021; Sawadogo & Darmont, 2021; Chen 2022):

- La **capa d'ingesta** (*ingestion layer*). Aquesta capa gestiona la incorporació de conjunts de dades al *data lake*, i inclou les funcions d'extracció de metadades (o metadescripció dels *datasets*) i modelat.
- La **capa de manteniment** (*maintenance layer*). Aquesta capa proveeix funcions de preparació i organització dels *datasets*, descobriment de *datasets* relacionats, interaccions entre dades, enriquiment de metadades, millora de la qualitat de les dades, i evolució de l'esquema.

- La **capa d'exploració** (*exploration layer*). Aquesta capa inclou les funcions de *query-driven data discovery* i consulta de dades heterogènies.

No hi ha tanmateix un consens general sobre l'estructuració en tres capes d'un *data lake* (Chen 2022). S'argumenta que, atès que un *data lake* es basa en la inexistència d'esquemes predefinits, no té sentit el modelat de dades. Wieder i Nolte, en la seva anàlisi del concepte d'arquitectura del *data lake*, descriuen diferents models per estructurar els components d'un *data lake*, que inclouen “[...] *the infrastructure, data storage, data flow, data modeling, data organization, data processes, metadata management, data security and privacy, and data quality.*” (Giebler et al., 2021).

L'arquitectura per zones (Zikopoulos et al., 2015) classifica les dades en zones en funció del seu nivell de processament, i esmenta un *raw data zone* per a les dades en brut (la capa d'ingesta), una zona per a les dades pre-processades, que anomena *staging zone* o *processing zone* (la capa de manteniment), i una capa per a la governança anomenada *trusted zone* o *sensitive zone* (la capa d'exploració) (Zikopoulos et al., 2015; Wieder & Nolte, 2022). L'arquitectura anomenada *Lambda* (Marz & Warren, 2015) busca compatibilitzar la conservació en el *data lake* de les dades en brut amb una acceleració en la capacitat de processament, que aconsegueix paral·lelitzant els processos de tractament en dues capes (*speed layer* i *batch layer*). D'altres arquitectures combinen el *data lake* amb un *data warehouse*, com el model *data lakehouse* (Armbrust et al., 2021), o encapsulen les dades per facilitar l'explotació, com es fa en la *FAIR digital object-based architecture* (Wieder & Nolte, 2022).

Però malgrat això, dins un *data lake* es troben casos en els quals cal parlar de models i modelat de dades, especialment pel que fa al modelat de metadades. Les metadades són un element fonamental per ordenar la informació dins un *data lake*, a través de la caracterització dels diversos i heterogenis *datasets* emmagatzemats en el *data lake*. De fet, són el primer àmbit de treball amb les dades, o la primera operació que es realitza sobre les dades en brut (també anomenada *dataset discovery* o *data profiling*), i la que vincula les dades amb les necessitats d'informació dels usuaris (Abedjan, Golab, i Naumann 2016, Chen 2022). En alguns casos s'ha experimentat amb l'ús d'ontologies per optimitzar el procés de *data profiling* (Alserafi et al. 2020).

L'ús d'un model distribuït no invalida l'ús complementari d'altres tecnologies per a la consolidació de les dades. Ans al contrari, alguns autors ho recomanen. És el cas de l'MDM, el qual té la mateixa utilitat en un *data lake* que en qualsevol altre entorn d'emmagatzematge de dades (eliminar inconsistències, resoldre duplicitats i crear una versió única i certa de cada dada, o *golden record*). La seva implementació es pot fer d'acord amb dos models (Laurent et al., 2019):

- MDM extern al *data lake*. Un sistema extern, a mode de *hub* amb el qual interaccionen totes les entitats del *data lake*. Es construeix de la mateixa manera que qualsevol MDM tradicional. I doni quan el *data lake* se suma a una arquitectura de sistemes preexistent, on ja es disposa d'un MDM.

- MDM intern al *data lake*. S'aplica en organitzacions amb un nombre molt alt de dades. Es construeix navegant pel *data lake* i aplicant sistemes d'analítica avançada de dades per part de científics de dades. Les dades mestres, un cop validades, són emmagatzemades en les pròpies taules del *data lake*. I doni quan el *data lake* és el centre de l'arquitectura de sistemes i emmagatzemament de dades de l'organització, i no es disposa prèviament de cap MDM.

La identificació d'informació rellevant dins un *data lake* és una de les àrees amb més projecció. A diferència d'un *data warehouse*, en un *data lake* no és necessari assolir cap estadi final de desenvolupament per a començar a fer explotació eficaç de la informació. El descobriment d'informació rellevant dins el *data lake* esdevé doncs un procés continu, on tècniques basades en *machine learning* i la identificació de patrons de coneixement ofereixen el seu major potencial (Bagozi et al. 2019; Papastefanatos, Meimaris, i Vassiliadis 2022). La utilització d'algoritmes de *machine learning* i IA i la contextualització progressiva dels continguts d'un *data lake* per optimitzar l'analítica de dades han portat a alguns autors a encunyar el concepte de *knowledge lake*, i també el d'*intelligent knowledge lake* (Beheshti et al. 2018; Beheshti, Benatallah, et al. 2020; Beheshti, Yakhchi, et al. 2020).

Una de les tasques cabdals de la integració de dades (en un *data lake*) és la *entity resolution*, que consisteix en encaixar (*matching*) diferents representacions d'entitats provinents de fonts diverses. Aquest procés utilitza tècniques de *blocking* (construcció de blocs per entitats a partir de l'anàlisi de conglomerats d'entitats similars) i de *meta-blocking* (reestructuració dels blocs anteriors en nous blocs per eliminar redundàncies). Aquesta tècnica (anomenada *query driven entity resolution*) ha estat especialment estudiada per l'equip de George Papadakis (Papadakis, Papastefanatos, i Koutrika 2014; Papadakis et al. 2015; Efthymiou et al. 2017).

El model *data lake* manca encara, segons alguns autors, d'una formulació sistemàtica i d'un marc de recerca consolidat (Chen 2022). També és qüestionat el seu rendiment quan manca d'un sistema fort de governança de la informació que eviti que es converteixi en un magatzem caòtic de continguts (Reina 2023).

2.2.3.5 L'evolució del *data lake*: el *data fabric* i la virtualització

El model *data lake*, tot i la flexibilitat de la seva arquitectura distribuïda pel que fa al manteniment i explotació, requereix igualment un procés de captura i replicació de les dades dins el repositori. Aquest procés, que gestiona la capa d'ingesta, implica no només la connexió per diferents canals a múltiples fonts de dades heterogènies, sinó la capacitat de detectar els canvis en les fonts originals per poder-ne activar la replicació. Això fa que l'actualització de les dades en un *data lake* esdevingui un repte major fins i tot que en els sistemes basats en ETL tradicionals (Stamper & Cervellati, 2023).

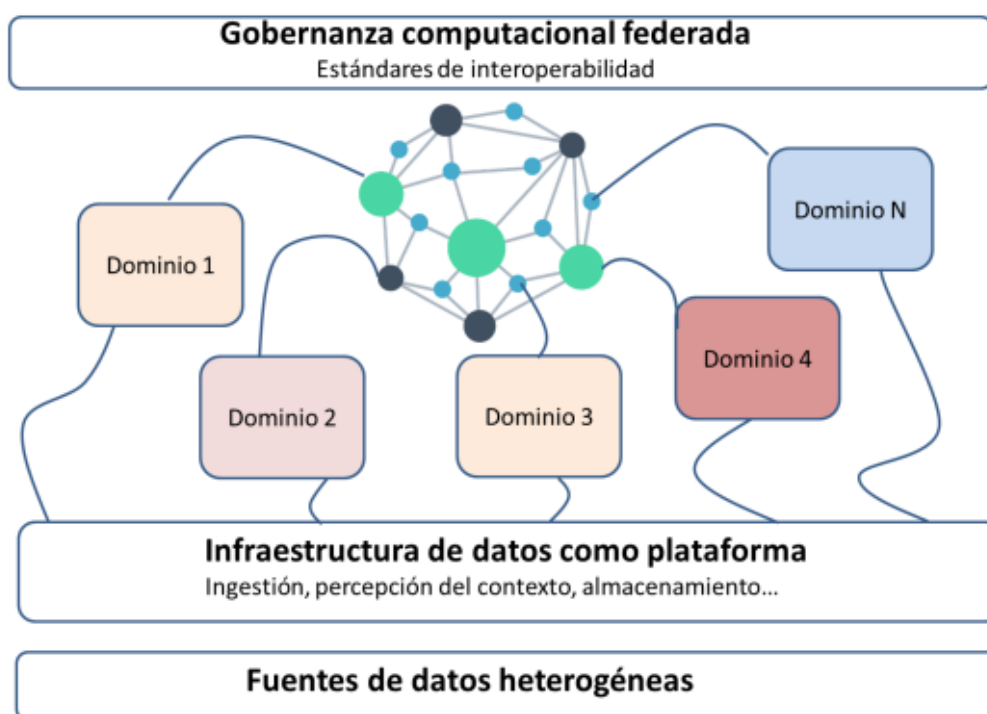
Diferents enquestes realitzades a científics de dades (Johnson & Dasu, 2003; Kandel et al., 2011; Lemieux et al., 2014) han posat en relleu quins són els principals problemes amb què es troben en realitzar projectes d'analítica avançada. Un d'ells és "descobrir" i localitzar on es troben les dades, conèixer l'existència i la ubicació de les dades, a causa de la manca d'inventaris de dades o d'inventaris deficients de sistemes d'informació. Un altre és la fragmentació de les dades en diferents ubicacions, i la dificultat per poder-hi accedir, sigui per la manca de coneixement de l'organització i de la seva estructura funcional, sigui per barreres de tipus legal o drets d'ús associats al context dels dades. I un tercer és la manca de fiabilitat de les dades, tant pel que fa a la seva veracitat com a la seva autenticitat, o el desconeixement de les condicions de generació, del context funcional i orgànic de la captura, o de la cadena de custòdia. Un model *data lake* pur no dona resposta, doncs, a aquestes necessitats.

Però l'esperit d'un *data lake* és precisament aquest, emmagatzemar les dades en brut fins que l'analista les necessiti i decideixi depurar-les (Beheshti et al., 2020). Per aquest motiu han proliferat els esforços per automatitzar el coneixement i tractament dels *datasets*, juntament amb la seva metadescrípció (Alserafi et al., 2020). Aquesta contextualització dels *datasets* és el que porta alguns autors a proposar el terme *knowledge lake*, al qual ja hem fet referència, i quan per aquesta contextualització s'utilitzen tècniques d'intel·ligència artificial, de *intelligent knowledge lake* (Beheshti et al., 2020).

Un *data fabric* es defineix com un sistema que proporciona una arquitectura unificada per gestionar i explotar dades (Östberg et al., 2022). Habitualment es tracta de sistemes distribuïts orientats a serveis, que es focalitzen en la gestió de dades i la connexió entre sistemes, sense tenir en principi una orientació de servei a l'usuari final. Poden connectar-se tant a recursos *on premise* com a fonts de dades al núvol, i actuen com un integrador d'informació de fonts diverses.

El model *data fabric* s'ubica tanmateix en el marc del *data discovery* (Rajmohan et al., 2023), i esdevé l'entorn idoni per a desplegar un sistema integral de governança de la dada. De fet, un *data fabric* és essencialment un element de governança, que permet controlar les dades en múltiples ubicacions i models (Moon et al., 2021). I és a través d'aquest context que també pot oferir interfícies per a la cerca i anàlisi de les dades, com en el cas de eines de virtualització (Östberg et al., 2022).

Figura 2-1: Estructura conceptual del model data fabric (Font: Delgado, 2022)



Aquesta capa d'exploració és la que permet, d'una forma controlada, “democratitzar” de forma efectiva l'accés a les dades curades segons un model federat o *data mesh* (Delgado, 2022). La clau és el tractament a nivell descriptiu que en un model *data fabric* es fa de les fonts de dades i de les dades que contenen, el qual sovint utilitza tecnologies semàntiques i knowledge graph (Schmid et al., 2019; Li et al., 20220). Per aquest motiu Ana-Maria Ghiran i Robert Andrei Buchmann afirmen que (Ghiran & Buchmann, 2019):

“The notion of “data fabric” is bridging the traditional gap between data and knowledge, by resorting to knowledge-based integration techniques [...] in order to bring together, in an operational way, heterogeneous data that spans across a multitude of sources or formats.”

Un dels elements que, tot i existir en el model *data lake*, cobra especial importància en el model *data fabric* és el *business glossary* (Sarkar, 2015). El *business glossary* és un element de negoci, no tecnològic, i té per funció descriure les dades des del punt de vista de la seva significació, normalitzant la seva denominació, fixant el seu responsable (*data steward*), i assignant-les a una àrea o domini de negoci (Sarkar, 2015; Kononow, 2019). Aquest és sens dubte un dels elements més importants per establir el punt de contacte entre aquest model tecnològic, les dades que conté i els instruments arxivístics, i és el que ha inspirat en aquesta tesi l'entitat “dada elemental” del model de dades i el procediment de mapatge de dades.

2.2.3.6 Els models orientats a objectes: present i futur dels gestors documentals

Als diferents models descrits per a la gestió d'informació estructurada cal afegir finalment la solució per excel·lència de la gestió documental: els sistemes de gestió electrònica de documents (EDMS). Les característiques d'aquests sistemes han estat sobradament descrites i formen part el corpus normatiu de l'arxivística i la gestió documental, especialment amb la segona versió de la ISO 15489 i les tres parts de la norma ISO 16175,⁷⁴ juntament amb les diferents versions de MoReq.⁷⁵

Els gestors documentals es basen majoritàriament en dos components:

- Un sistema de fitxers, que conté els objectes digitals.
- Una base de dades orientada a objectes, que conté les metadades dels objectes i els apuntadors als fitxers que conformen cada objecte.

Els gestors documentals (EDMS) han estat les eines amb les quals els arxivers s'han sentit més còmodes, i s'han considerat en l'àmbit arxivístic la pedra angular de qualsevol sistema de gestió de documents electrònics, esdevenint el sistema de referència del model docucèntric. Però al mateix temps han consumit l'energia de bona part de la teoria arxivística formulada durant el darrer decenni, i de la pràctica en la construcció d'estructures massives de carpetes (Morelli, 2007; Jones, 2008) i de complexes regles de classificació (Ludlow, 2016).⁷⁶

La “descoberta” per part de l'arxivística dels gestors documentals es remunta a la dècada dels 90, quan els documents d'oficina es començaven a multiplicar sense control a les xarxes locals. Fou en aquest moment quan es va generar la idea que un EDMS era peça indispensable en un sistema de gestió dels documents electrònics, que havia d'esdevenir la “*single source of truth*” de l'organització (Ludlow, 2016). El primer model de desplegament consistia en una plataforma única per a tota l'organització, connectada quan era possible als sistemes de gestió, i en la qual els usuaris dipositaven els documents de forma eminentment manual. Aquesta plataforma funcionava al mateix

⁷⁴ ISO 16175-1:2010 – *Information and documentation – Principles and functional requirements for records in electronic office environments – Part 1: Overview and statement of principles*, 2010; ISO 16175-2:2011 – *Information and documentation – Principles and functional requirements for records in electronic office environments – Part 2: Guidelines and functional requirements for digital records management systems*, 2011; ISO 16175-3:2010 – *Information and documentation – Principles and functional requirements for records in electronic office environments – Part 3: Guidelines and functional requirements for records in business systems*, 2010).

⁷⁵ Versió vigent i anteriors accessibles a <https://www.moreq.info/specification>.

⁷⁶ El debat sobre la creació de carpetes (*folder*) o l'ús de metadades per crear les agrupacions documentals en els EDMS és un exemple de la forma com aquesta tecnologia ha centrat el debat arxivístic durant aquest temps.

temps com arxiu semiactiu digital, i com a infraestructura d'emmagatzematge de fitxers per a certes aplicacions de negoci. Progressivament se li afegiren components per a gestionar les polítiques de retenció i la limitació de l'accés en fase d'arxiu, en el que s'anomenava *Electronic Records Management Systems* (ERMS) (Kampffmeyer, 2023).

Quan els arxivers van adoptar els gestors documentals com una eina pròpia, van començar a prendre forma tres models per a la gestió dels documents electrònics. El primer model, representat per l'escola canadenca, consistia en reproduir la gestió basada en paper, i realitzar periòdicament transferències documentals, en aquest cas digitals, entre els sistemes generadors dels documents i una plataforma transversal, actuant a mode de dipòsit digital semiactiu, on la documentació s'estructurava segons criteris arxivístics (Duranti & MacNeil, 1996; Duranti et al., 2002). Aquest model, anomenat per James Lappin "*Separate records repository model*" (Lappin et al., 2021), ha estat el més utilitzat per part de la gestió documental a Europa, i especialment a Catalunya, i encara marca la pauta de les arquitectures de gestió documental del sector públic català.⁷⁷ Aquest model va inspirar l'estàndard DoD 5015.2 del Departament de Defensa dels Estats Units, i es va recollir posteriorment en la primera versió dels *Model Requirements for the Management of Electronic Records* (MoReq) i els sis llibres blancs de la indústria del DLM-Forum (Serra, 2002).

El segon model fou el model australià, consistent en intervenir directament en el disseny de les aplicacions de negoci de les organitzacions per tal de garantir que incorporin funcionalitats de gestió documental (Bearman, 1991, 1994). Aquest model, que James Lappin anomena "*Intervention in business applications model*" (Lappin et al., 2021), ha inspirat la gestió documental electrònica al món anglosaxó, i ha determinat en bona part la configuració de la segona versió de la norma ISO 15489, i la família de normes ISO 16175, com també la tercera versió de MoReq (MoReq2010).

I en els últims anys ha agafat força un tercer model, el que s'ha anomenat "*records management in-place*", que consisteix en conservar els documents en els sistemes origen, malgrat que aquests sistemes no disposin de funcionalitats específiques de gestió documental, i gestionar-los mitjançant regles externes que poden actuar dins de cada sistema (Lappin et al., 2021). Aquest model es vincula al corrent que cerca la simplificació de les tasques de gestió documental, i en són exemples el model Capstone del NARA (*White Paper on The Capstone Approach and Capstone GRS*, 2015), i la recent evolució dels entorns de col·laboració al núvol (Lappin, 2010, 2020, 2021; Sarkar, 2015). Alguns consultors recomanen aquest model quan la simplicitat dels processos a gestionar, el volum de documents o el nombre de carpetes que cal crear dins el "*file plan*" no justifiquen el desenvolupament d'un EDMS ni la modificació d'aplicacions ja existents (Caplinger & Borgschulte, 2015).

⁷⁷ Vegeu el model Aresta a <https://cultura.gencat.cat/ca/temes/arxius-i-gestio-documental/gestio-documental/gdgencaat/aresta/index.html>.

De fet, la complexitat de connectar tots els sistemes d'informació d'una organització a una única plataforma transversal (el "*Separate records repository model*") s'ha considerat una de les causes de la baixa implantació de sistemes de gestió documental tipus EDMS a les organitzacions.⁷⁸ És per aquest motiu que Cassie Findlay afirmava en un article de referència que "*We should stop talking about: 'putting things in' systems*" (Findlay et al., 2014). Tant el segon com el tercer model es troben també en la línia del post-custodialisme d'arrel australiana (McKemmish et al., 2019), hereu del model del *continuum* proposat per Frank Upward (Upward, 1996, 1997, 2000, 2019).

Des del principi els gestors documentals oferien la possibilitat de gestionar un tipus particular d'objecte, compost només de dades i sense cap vinculació necessària amb cap fitxer (un objecte buit).⁷⁹ A mesura que la gestió de les metadades evolucionava dins els EDMS (incorporació de taxonomies, personalització de metadades, vinculació a bases de dades externes via ODBC, etc.) s'obrien noves possibilitats de gestió d'informació estructurada utilitzant eines EDMS. Això va portar amb el temps a l'aparició dins els EDMS de nous tipus d'objecte orientats clarament a la gestió d'informació estructurada, com les llistes de *Microsoft Sharepoint*. Sharepoint incorporava funcionalitats que posaven a l'abast de l'usuari final la creació de llistes (en essència taules de dades), i també la gestió de les metadades de qualsevol objecte de forma autònoma i amb grans possibilitats d'explotació. Aquestes possibilitats foren tanmateix poc explotades per les implementacions d'aquestes plataformes, especialment les liderades, llevat d'algunes excepcions, des de l'àmbit arxivístic (Serra, 2006).

Malgrat moltes organitzacions encara continuen en la fase d'una única plataforma centralitzada de gestió documental, si és que l'han assolit, el model va evolucionar amb noves estratègies per a la gestió documental electrònica (Lappin, 2021). Progressivament es va descentralitzar la gestió documental, traslladant-la als sistemes de gestió o negoci, i incorporant en aquests sistemes funcionalitats de gestió documental. Així van aparèixer solucions dedicades, com els *e-Mail Management Systems* o el *Digital Assets Management Systems* (Kampffmeyer, 2023). L'evolució va continuar dotant progressivament als entorns de negoci de funcionalitats de gestió

⁷⁸ Cassie Findlay cita els resultats d'una consultoria realitzada per James Lappin, on, en una organització que disposava d'un potent sistema EDMS implantat, només el 4% del volum total d'informació de l'organització estava custodiat en aquesta plataforma. Del 96% restant, un 33% estava contingut en aplicacions de gestió (majoritàriament dades estructurades), i un 63% estava contingut a les bústies de correu electrònic i a la xarxa local (Findlay et al., 2014).

⁷⁹ L'evolució en la gestió de les metadades dels objectes cada cop ha anat agafant més protagonisme fins a convertir-se en una de les fonts d'informació més rellevants del gestor documental. Aquesta funcionalitat no deixa de ser l'evolució del tradicional ús que el personal gestor o administratiu han fet dels fulls de càlcul per a generar registres d'activitat administrativa, que amb el temps han esdevingut bases de dades domèstiques.

documental, fent irrellevant la transferència digital a un sistema centralitzat, en la línia del ja esmentat *records management "in place"* (Lappin, 2021).⁸⁰

Però durant la segona dècada del segle XXI dos fets van provocar un canvi substancial en aquests plantejaments (Oladejo & Hadžidedić, 2021):

- El primer fet és la disponibilitat dels gestors documentals en modalitat *SaaS*.
- I la segona és la seva comercialització en forma de suites que integren diferents eines de productivitat o entorns de treball virtuals.

L'entorn que va obrir la porta a aquest model fou la suite de *Google*,⁸¹ que va superar els gestors documentals *on premise* en aquell moment líders del mercat en agilitat i flexibilitat. Però la revolució ha vingut de la mà de la posada al núvol de Microsoft Sharepoint i de la seva integració amb la suite *Microsoft 365*.⁸² L'aparició de *Microsoft TEAMS* ha propiciat un desplaçament del focus d'atenció dels treballadors, passant de la gestió d'objectes d'informació (correu electrònic i documents en xarxa) a la gestió d'un flux continu d'informació (xats i canals de TEAMS). Els documents, compartits a través de xats i reunions en línia, i editables en línia de forma concurrent, estan passant a formar part d'aquest flux d'informació, amb la qual cosa la fixació d'un contingut estable que es pugui arxivar segons els paràmetres tradicionals es fa cada cop més difícil (Foscarini et al., 2022).

Microsoft TEAMS és una eina de comunicació, basada en la missatgeria. Invalida per tant la possibilitat de mantenir l'estratègia d'una definició apriorística de les estructures d'informació, el que James Lappin anomena "*records management by design*" (Lappin, 2021). Lappin alerta que es corre el risc de tornar a models anteriors a la intervenció de la gestió documental en la creació dels documents, quan s'actuava a posteriori intentant reconstruir l'ordre original, i de fet alguns autors ja parlen en l'entorn de TEAMS, però també en altres entorns digitals (com la xarxa local) de la necessitat d'una "*retrospective governance*" (Sherman, 2021).

La culminació d'aquesta evolució està arribant amb la introducció de la intel·ligència artificial en els entorns de treball col·laboratiu actuals (Lappin, 2020a). Les dues grans plataformes tenen en el seu full de ruta posar en producció aquests serveis abans d'acabar l'any 2023, amb un lleuger avantatge de Microsoft sobre Google. Aquesta avantatge prové del projecte que Microsoft va presentar durant la trobada *Ignite* 2019, el projecte *Cortex* (Patton, 2019; Lappin, 2020c). Els tres anys d'experiència després del

⁸⁰ Llevat de la preservació a llarg termini en sistemes basats en el model OAIS, que sempre s'ha plantejat com un entorn totalment a part dels entorns de negoci (Richards, 2016; Wilson, 2017).

⁸¹ Anomenada inicialment *G-Suite*, i actualment *Google Workspace* (<https://workspace.google.com/>).

⁸² Anomenada inicialment *Microsoft 365*, i combinada després amb altres mòduls (*Power Tools*, *Microsoft TEAMS*, etc.), fins convertir-se en l'actual suite *Microsoft 365* (<https://www.microsoft365.com/>).

llançament de Microsoft TEAMS van permetre observar un canvi substancial en el comportament dels usuaris, que obria la porta a tancar definitivament *Sharepoint* com a *front-end*, i eliminar de les tasques d'usuari final qualsevol acció de classificació, declaració de document d'arxiu, assignació de metadades o recuperació. El projecte inicial contemplava tres components:

- *Microsoft Graph*, com un entorn en forma de xarxa a través del qual pot navegar l'usuari accedint a qualsevol actiu d'informació (document, contacte, etc.) o traça de fet que estigui contingut en Microsoft 365.⁸³
- *Microsoft Sharepoint*, com a back-end on s'emmagatzemaran tant dades com documents, i on es farà tota la gestió de metadades i cicle de vida *in place*.
- *Microsoft AI*, un motor basat en intel·ligència artificial que serà l'encarregat de recollir, classificar, metadescriure i etiquetar tots els actius, i oferir-los a través del canal més adient en funció de les necessitats de l'usuari (xat o canal de TEAMS, correu electrònic, etc.).

La interfície de treball per a l'usuari final havia de rebre el nom de *Knowledge center*.

Dos elements de gran impacte sobre la gestió documental del projecte Cortex són:

- Qualsevol element sobre el qual es puguin recollir dades passa a ser un actiu d'informació en una mateixa xarxa de coneixement. Es trenca doncs la frontera entre documents i agrupacions documentals, i la resta d'elements dadificables (usuaris i persones de contacte, aplicacions, flux de treball, etc.).
- Això té una derivada immediata: la classificació ha de permetre classificar qualsevol mena d'actius, no només documents. Per aquest motiu, el projecte Cortex plantejava una classificació basada en *topics*, és a dir temàtica, i generada de forma orgànica i partint dels usos que en fessin els usuaris, que el motor de IA analitzaria. Aquests *topics* serien accessibles de forma indistinta a través de qualsevol aplicació d'Microsoft 365 (Outlook, Excel, TEAMS, Automate, etc.).

El projecte Cortex preveia a més la possibilitat d'utilitzar IA generativa per a la creació del que anomenava *Topic pages*, és a dir pàgines que agregaven tota aquella informació i serveis relacionats amb un *topic* o necessaris per a la seva gestió, que podrien ser també editables per l'usuari. En tractar-se d'un motor d'IA, l'ús i edició que fes l'usuari de cada *topic* seria utilitzat per refinar la categorització de tots els actius del sistema.

Per a fer possible una gestió eficaç de les metadades administrades, fins ara integrades en Sharepoint, el projecte Cortex preveia segregar aquesta funcionalitat en un mòdul transversal, el *Managed Metadata Service* (MMS), accessible des de qualsevol component de Microsoft 365.

⁸³ Aquest component es basa en la idea de *knowledge graph*, element essencial en la construcció d'ontologies i molt rellevant per al funcionament de motors d'intel·ligència artificial.

L'any 2020 Microsoft va canviar el nom del projecte Cortex pel de Viva Topics, integrat dins el nou component de Microsoft 365 anomenat Microsoft Viva, presentat com una experiència integral de treballador. Viva Topics⁸⁴ incorpora doncs la part d'experiència d'usuari (Topic pages, integracions amb Outlook i TEAMS, etc.), mentre que el motor d'AI responsable de l'anàlisi i etiquetat queda integrat amb Sharepoint amb el nom de Microsoft Syntex.⁸⁵ El 16 de març de 2023 Microsoft va anunciar la darrera incorporació a Microsoft 365: Copilot (Spataro, 2023; Pires, 2023). Copilot és un assistent de productivitat basat en GPT4 de OpenAI, que permetrà crear continguts utilitzant la informació disponible a Microsoft 365 i especialment la facilitada per Microsoft Graph (Pires, 2023).

⁸⁴ Vegeu: <https://adoption.microsoft.com/en-us/viva/topics/>.

⁸⁵ Vegeu: <https://adoption.microsoft.com/en-us/syntex/adoption/>.

2.3 El constructe documental: la sèrie documental com a clau de volta de l'aportació arxivística

2.3.1 Característiques i evolució del concepte de sèrie documental a l'arxivística moderna

Una de les idees motriu de la gestió documental, des que aquesta va prendre forma amb Schellenberg (Schellenberg, 1956, 1965), és que el document, entès com unitat d'informació, no pot ser objecte de gestió de forma individual. L'argument que sosté aquesta afirmació és doble:

- En primer lloc, per un motiu operatiu: el nombre de documents a gestionar en qualsevol organització és tant elevat que la seva gestió només es abastable si s'agrupen.
- I en segon lloc, per un motiu conceptual: el valor del document es deriva del context funcional en el qual es crea i actua, un context que es materialitza en el context documental, i fora d'aquest context perd el seu valor (Duranti, 1997).

El fet que el valor del document es derivi del seu context, fa que la unitat de treball bàsica de la gestió documental no sigui el document sinó l'agrupació documental (Duranti, 1997, 1999; Yeo, 2012). Totes les metodologies de gestió documental es basen en agrupar els documents en conjunts significatius, els quals tenen en comú que són producte d'un determinat context (*DIRKS: A Strategic Approach to Managing Business Information*, 2001; Healy, 2001; *Alberta's modified functional classification system model*, 2007; Macintosh & Real, 2007).

La necessitat de gestionar els documents de forma agrupada no és una idea nova en l'arxivística, i des de finals del segle XIX ja formava part de les pautes de treball dels arxiviers en els processos de descripció dels fons d'arxiu (Jenkinson, 1922). I aquesta tradició no és absent del plantejament de la gestió documental. En tot cas, ha estat una constant que s'ha mantingut gairebé inalterable fins a principis del segle XXI (Buckland, 1994; Shepherd & Yeo, 2003; Yeo, 2009; Duff & Haskell, 2015; Mokhtar & Yusof, 2016).

Les formulacions d'aquesta necessitat són nombroses. Segons Elisabet Shepherd i Geoffrey Yeo (Shepherd & Yeo, 2003):

The record of a single step (or sub-step) may be of little significance in its own right. It acquires meaning when combined with the records of other steps, to form a record of the activity as a whole.

L'aparició de la gestió documental electrònica, cap a finals dels anys 90, no va modificar aquest plantejament, però sí que el va tensionar notablement (Bearman, 1991; *Committee on Electronic Records*, 1997; Suderman, 2002). L'increment de la capacitat de processament i recuperació d'informació que va aportar la tecnologia va començar a posar en qüestió la necessitat d'utilitzar estructures d'agrupació rígides o constants, i el debat sobre la creació o no de carpetes d'expedient dins un gestor documental va ser recurrent fins fa pocs anys (Frendo, 2007; Morelli, 2007; Rodríguez, 2012; Mokhtar et al., 2016).⁸⁶ Un exponent de la vivesa d'aquest debat el trobem a la norma MoReq 2010. La norma MoReq2010 especificava que els documents, en qualsevol àmbit de gestió documental, i preferentment en els entorns automatitzats, han de ser ubicats sempre dins una agrupació documental, i mai poden gestionar-se assignats a altres entitats del sistema que no siguin agrupacions documentals. Però al mateix temps qüestionava el vincle de les agrupacions documentals amb la classificació, i argumentava que la relació jeràrquica entre productor, activitat i agrupació documental (sèrie), tradicionalment utilitzada en els sistemes de classificació basats en paper (Rodríguez, 2012), era massa rígida i desaprofitava les possibilitats que ofereixen els gestors documentals. I posava com a exemple que moltes organitzacions agrupen els documents per projectes, que sovint corresponen a més d'una funció o activitat (MoReq2010, apartat 1.4.5).

Es començava doncs a posar en qüestió el primer motiu al·legat tradicionalment per a la creació de sèries: el volum de documents a gestionar. La capacitat de processament d'informació digital, ja no a nivell d'unitat documental, sinó a nivell de dada, ha fet abastable, a un cost cada cop més racional, la gestió i la descripció a nivell de document o evidència ("*item level*"), defensada des de fa una dècada per nombrosos autors (Bak, 2012; Yeo, 2013), confirmant la interpretació feta força més abans per David Bearman que les agrupacions documentals eren un "mal menor" que calia assumir per a fer possible la gestió de l'enorme volum de documents existent en paper (Bearman, 1991, 1994).

L'eclosió de la gestió basada en dades, i la possibilitat de gestionar mitjançant metadades múltiples facetes d'un mateix objecte, fins i tot fora d'un gestor documental (model *data lake*), ha accelerat aquesta progressiva erosió del concepte de sèrie documental, que, malgrat no l'ha fet desaparèixer dels models de descripció arxivístics,⁸⁷ sí que ho ha fet pràcticament de la literatura especialitzada.⁸⁸

⁸⁶ Aquest debat ha condicionat, per exemple, el desenvolupament de la política de gestió documental de la Generalitat de Catalunya (*Guia funcional del Gestor documental corporatiu (ARESTA)*, 2016), i un nombre important de desenvolupaments de solucions d'Administració electrònica que encara incorporen sistemes estrictes d'agrupació documental.

⁸⁷ Com es pot veure en el repàs dels models actuals consultable en el capítol 4. A títol d'exemple, el model *Records in Context* (RiC), el més actual, segueix contemplant el concepte d'agrupació documental en l'entitat "*Record set*".

⁸⁸ En la darrera dècada, en les principals revistes científiques de l'àmbit de coneixement de la gestió documental, pràcticament no es troben articles dedicats a fer evolucionar el concepte de

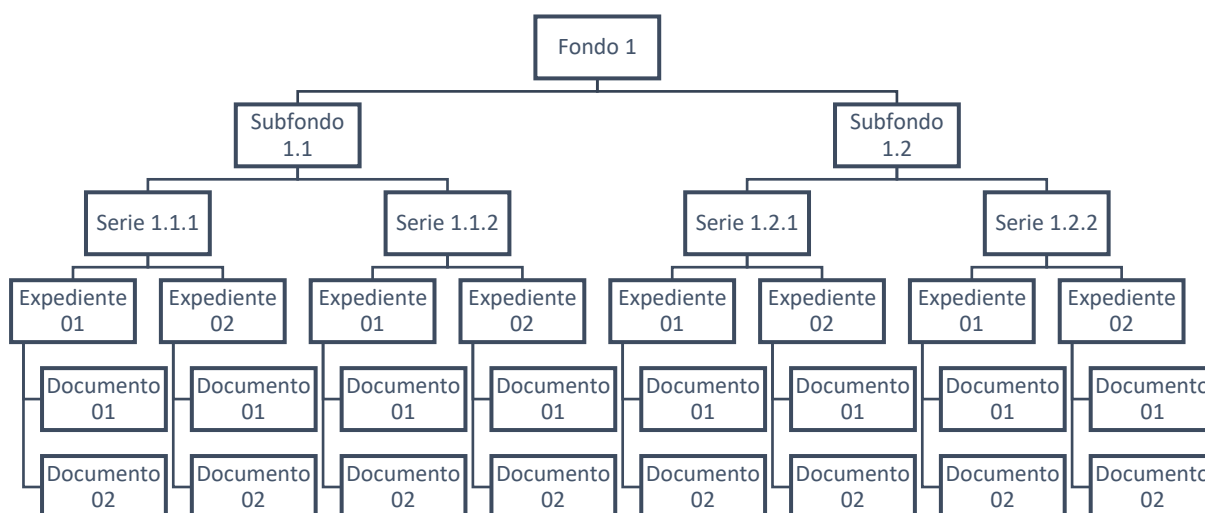
Revisem doncs com ha estat aquesta evolució del concepte de sèrie documental.

2.3.1.1 El concepte de sèrie documental i la seva estructura

La norma ISAD-G defineix sèrie documental com “*Conjunt de documents organitzats d'acord amb un sistema de classificació o conservats com a unitat pel fet de ser el resultat d'un mateix procés de formació o de classificació, o de la mateixa activitat, perquè tenen una mateixa tipologia, o per qualsevol altra relació derivada de la producció, recepció o ús. La sèrie també es coneix pel nom de sèrie documental.*” Aquest caràcter ampli del concepte de sèrie documental s'evidencia en l'afirmació de la consideració de la sèrie com “[...] *qualsevol altra relació derivada de la producció, recepció o ús*”.

La sèrie documental s'ha considerat tradicionalment un nivell concret dins una jerarquia d'agrupacions documentals (Yeo, 2009; Galarreta et al., 2012). El principal exponent on es condensa aquesta concepció és la norma ISAD(G), que identifica fins a cinc nivells (ampliables) en la jerarquia d'agrupacions documentals, diferenciant entre un primer nivell de caràcter orgànic (el fons i el subfons) i un segon nivell de caràcter funcional, on apareix la sèrie documental, la qual es descompon en unitats documentals compostes, i aquestes en unitats documentals simples.

Figura 2-2: Jerarquia d'agrupacions documentals ISAD(G)



sèrie o agrupació documental. Això no obstant, es troben gran quantitat d'articles descrivint contingut de sèries concretes, el que demostra la pervivència del concepte en la pràctica diària dels arxius, tant administratius com històrics.

Els models de descripció concebuts per ser aplicats tant en arxius administratius com històrics, com la norma ISAD(G), desgranen una tipologia d'agrupacions documentals amb valors molt similars, i incorporen tots ells el concepte de sèrie documental. Així doncs, la norma *Encoded Archival Description* (EAD) també contempla la sèrie documental de forma explícita, dins una extensa relació de nivells de descripció: *class, collection, file, fonds, item, otherlevel, recordgrp, series, subfonds, subgrp, subseries*. I també ho fa el model conceptual de las *Normas Españolas de Descripción Archivística* (NEDA), la qual contempla vuit tipus d'agrupació documental: *Grupo de fondos, Fondo, División de fondo o de grupo de fondo, Serie, Subserie, Fracción de serie/subserie, Colección* i *División de colección*.

No s'esdevé el mateix en les normes elaborades específicament per a la gestió documental. En aquest cas, en la majoria de casos es fa referència a un nivell d'agrupació documental genèric, sense utilitzar en principi una terminologia específica (Delgado, 2007). El model australià AGRkMS contempla l'entitat **record**, que defineix com: "[...] *A record may comprise an electronic or paper-based document or group of aggregated documents*". De la mateixa manera, MoReq2010 utilitza el concepte genèric de "aggregation" per representar qualsevol tipus d'agrupació documental. Només en el cas de les normes directament inspirades en la ISO 23081 es manté el concepte de sèrie, atès que la pròpia ISO ja el contempla com un dels sis possibles nivells de l'entitat "record". Així s'esdevé amb el model **e-EMGDE**, el qual contempla simultàniament, com dos possibles nivells de descripció (entre d'altres) els conceptes de "serie documental" i de "agregación".

La nota d'alerta apareix amb la norma més recent emanada de l'àmbit arxivístic, el model ontològic *Records in Context* (RiC). RiC no recull el concepte de sèrie documental a nivell d'entitat. Ho argumenta en l'apartat 1.6.2.1 RiC, afirmant que en la disciplina arxivística no existeix un acord unànime sobre el concepte de sèrie, atès que sovint s'ha treballat amb diferents nivells d'agrupació, no sempre homogenis ni en la concreció ni en el criteri d'agrupació (fons, subfons, col·leccions, sèries, subsèries, etc.), la qual cosa ha generat ambigüïtat en moltes descripcions arxivístiques.⁸⁹ Per aquest motiu, RiC proposa una única entitat neutra, que anomena "record set", la qual depèn de "record resource", tot i que a través de l'atribut "record set type" permet catalogar un *record set* com a sèrie documental.

RiC argumenta també que, atès que els diferents nivells d'agrupació, en el nivell funcional, comparteixen els mateixos atributs descriptius (com recull també la norma ISAD(G) a través de la descripció multinivell), es pot interpretar que són "the same kind of thing". Això no obstant, RiC desgrana l'entitat "record resource" en tres nivells de concreció: "record set", "record" i "record part" (RiC, apartat 2.1).

⁸⁹ En referència a la distinció entre document i agrupació, RiC afirma que "Archivists have long recognized that the two are not the same, but there has been no clear guidance on how to describe each, and this has led to inconsistent and ambiguous practice" (apartat 1.6.2.1).

En tots els casos doncs, incloent RiC, les agrupacions documentals es conceben com estructures jerarquitzaes, que s'agrupen a diferents nivells, amb l'objectiu d'heretar propietats i comportaments dels nivells superiors, especialment en l'entorn digital (el que MoReq2010, en el punt 6.2.2, descriu com a “*root aggregations*”).⁹⁰

2.3.1.2 La visió utilitarista de la sèrie documental: la sèrie com a recurs per a la gestió

La importància cabdal que el concepte de sèrie, i per extensió d'agrupació documental, té per a la gestió documental, és la capacitat de provocar efectes sobre els documents o *records* (Convery, 2014; Cumming & Picot, 2014; Anderson et al., 2015). A diferència de l'arxivística clàssica, la gestió documental té un caràcter actiu, que pot esdevenir fins i tot transformador del funcionament de les organitzacions (Millar, 2017). I aquest caràcter es desplega a través del concepte de sèrie. Per això és tant important aquest concepte per a la disciplina arxivística (Guardado, 2013).

Sense restar rellevància al paper de la sèrie documental com a entitat documental amb significació pròpia dins un sistema de descripció arxivística, i com a recurs per a la recuperació d'informació, la gestió documental centra l'atenció en l'aspecte operatiu de la sèrie documental, i la considera un **recurs per a la gestió**, que permet determinar:

- La forma i estructura dels documents que la componen.
- El valor d'aquests documents i els sistemes d'autenticació a aplicar.
- El destí intermedi i final de cada document, i el format en què haurà de perviure.
- El règim d'accés, utilització i publicitat de cada evidència al llarg del temps.

Això és el que fa especialment potent el concepte de sèrie documental, ja que actua com a recurs reductor d'una realitat voluminosa i formalment diversa (Roberge, 1992). Permet relacionar documents de formes diverses com pertanyents a la mateixa realitat funcional, vincular-los al seu context, i programar fets determinants del seu cicle de vida (Duranti, 1997).⁹¹

Aquesta visió utilitarista es especialment present en l'arxivística anglosaxona, i complementa la definició de la sèrie documental, afegint a la connexió amb el procés

⁹⁰ Tot i això, MoReq2010 recomanava utilitzar el mínim possible de nivells en les agrupacions, per tal d'evitar acabar generant pseudo-classificacions jeràrquiques. I en tot cas no permetia arxivar documents producte d'un mateix procés o activitat a diferents nivells de la jerarquia d'agrupacions, atès que “*Placing records at different levels of aggregation breaks the natural linear sequence of records within the aggregation and thus interrupts its narrative.*” (punt 6.2.7).

⁹¹ Aquesta afirmació es compleix en el cas d'aplicació d'un criteri de classificació funcional. Les classificacions per tipologies documentals, freqüents en l'Europa continental, no fan possible l'abstracció de la significació del document de la seva forma externa (García Ruipérez, 2007).

generador el fet que els elements que componen la sèrie tinguin unes característiques comunes (Yeo, 2012)

De fet, l'objectiu primer d'un sistema de gestió documental és vetllar per la conservació i un accés i usos adequats dels documents d'una organització (Harries, 2009; Wright, 2013), tal com estableix la norma ISO 15489 i els models que s'hi inspiren (MoReq, e-EMGDE, etc.). Cal destacar que la norma ISO 15489, com a components d'un sistema de gestió documental, a part del quadre de classificació funcional cita dos elements més: el quadre de seguretat i accés, i les regles de conservació i eliminació.

Així doncs, l'assignació d'una unitat documental (simple o composta) a una sèrie documental no té altre objectiu que fer possible l'aplicació sobre aquesta unitat documental d'una sèrie de regles de comportament dins el marc d'un sistema de gestió documental (Joyce, 2002; Shepherd & Yeo, 2003; Cisco, 2008a; Foscarini, 2012). Malgrat les diferents propostes i models, totes les regles tenen una mateixa estructura d'execució (Craig, 2004; Vázquez Murillo, 2006; Vieira et al., 2012):

- Un **esdeveniment** que dispara l'execució de la regla (*trigger*). L'esdeveniment pot tenir caràcter documental (per exemple el tancament d'un expedient) o estar vinculat al context d'execució del procés que dona origen a la sèrie documental (per exemple un esdeveniment fortuït, un accident, o un canvi normatiu, etc.). Aquest esdeveniment, en cas que la regla s'hagi d'executar automàticament, ha de poder ser detectat també de forma automàtica pel sistema de gestió documental.
- Un **termini** a comptar a partir de l'esdeveniment. Tot i que l'execució d'una regla pot ser immediata a l'esdeveniment, el més habitual és preveure un termini de caràcter precaucional per evitar que l'execució de la regla pugui tenir cap conseqüència no prevista o negativa. Aquest termini es pot fixar i comptar en qualsevol unitat temporal.
- Una **resolució**, entesa com acció que modificarà l'essència o l'existència de la unitat documental. Aquesta resolució pot implicar l'eliminació (total o parcial) de la unitat documental, un canvi de responsabilitat en la seva custòdia, una transformació o canvi de format, una modificació dels permisos d'accés (usuaris que hi poden accedir), i una modificació dels permisos d'ús. La norma MoReq2010, per exemple, parla de resolucions de destrucció, transferència, exportació i revisió.

El principal efecte de l'aplicació d'una regla sobre una unitat documental és que provoca un **canvi en l'estat** de la unitat documental, un canvi que, a més d'afectar en alguns casos les seves característiques intrínseques, condiona el **tractament** que rebrà la unitat documental mentre es mantingui en aquell estat (on es conservarà, com es descriurà, qui hi podrà accedir, etc.). El temps que la unitat documental romandrà en aquell estat és el que anomenarem **fase**, i el conjunt de fases o estats per les quals

passa una unitat documental rebrà el nom de **cicle de vida** (Craig, 2004; Vázquez Murillo, 2006).

Però la sèrie documental ha anat adquirint amb el temps nous usos. El concepte de sèrie documental esdevé un element indispensable per a l'anàlisi i millora de processos (Chandler, 2022; Escobar, 2009; Niu, 2014). Quan s'analitza el circuit documental d'una tramitació, abans de plantejar propostes de millora cal conèixer amb detall el substrat documental vigent, la composició, l'abast i les deficiències. En la majoria dels casos, són diverses les sèries documentals que es veuen afectades per cada procés (Convery, 2014; Cumming & Picot, 2014). L'anàlisi combinat d'aquestes sèries permet simular l'impacte que el redisseny d'un procés pot suposar sobre les evidències i sistemes d'informació d'un procés, i calcular el risc i els beneficis d'implementar determinades accions de millora (Serra & Casellas, 2007; Serra, 2006).

2.3.1.3 El caràcter interseccional de la sèrie documental

Una dècada després de Schellenberg, l'arxiver australià Peter Scott va fer una aportació cabdal per a definir el concepte de sèrie documental (Scott, 1966). El sistema de classificació definit per Schellenberg feia coincidir en un mateix instrument la descripció dels documents amb la descripció del seu context funcional, d'acord amb una estricta estructura en tres nivells (funció, activitat i transacció). I malgrat el seu caràcter funcional, no descartava la possibilitat d'utilitzar el criteri orgànic en aquelles organitzacions on l'estructura orgànica reflectís clarament l'estructura funcional. Scott, observant la creixent mobilitat tant en les estructures organitzacionals com en l'assignació de funcions a òrgans, va proposar un sistema on se separaven la descripció dels documents de la descripció del context, un context que es descomponia en dues vessants:⁹²

- Els agents que creen i reben els documents
- Els processos (funcions i activitats) que donen origen als documents

D'aquesta manera, la sèrie documental esdevenia allò que es crea quan les persones i les organitzacions realitzen funcions i activitats (Cunningham et al., 2013). Aquesta concepció de la sèrie documental, basada en el triangle entre agents ("*people*"), processos o activitats ("*business activities*") i documents ("*records*"), va conformar el que s'ha anomenat l'"*Australian series system*" (Feldman, 2003; Cunningham et al., 2013; Hurley, 2008).

Però la influència d'aquest model es va multiplicar quan aquest triangle d'entitats va constituir el nucli de la norma australiana de descripció arxivística, l'*Australian*

⁹² De fet, separava les dues dimensions que en la norma ISAD(G) es formulen en forma jerarquitzada: la dimensió orgànica i la dimensió funcional.

Recordkeeping Metadata Schema (McKemmish et al., 1999; Acland, 2006), precedent de l'*Australian Government Recordkeeping Metadata Standard (AGRkMS)*. Aquesta norma fou adoptada com a base de la norma ISO 23081, de recomanada utilització per la norma ISO 15489, amb la qual cosa es va convertir en l'estàndard per a la descripció en gestió documental (Cottin & Dessolin, 2012).

El sistema de sèries australià també va servir de base a la teoria del *continuum* (Upward, 1996, 1997, 2000, 2019; McKemmish, 1997; Upward et al., 2013), la qual va marcar un punt d'inflexió en la teoria majoritària del cicle de vida enunciada per Schellenberg (Schellenberg, 1956; Vázquez Murillo, 1995, 2006).

És en aquesta concepció interseccional de la sèrie documental, estandarditzada per la norma ISO 23081, que hem basat la metodologia d'identificació funcional desplegada en els capítols 3 i 4 d'aquesta tesi. Ho hem plantejat partint de les dues dimensions ja exposades:

- La **dimensió funcional**, la qual prové de la relació que el concepte de sèrie manté amb una determinada funció o activitat de l'organització. De fet, podríem definir la sèrie documental com una "funció o activitat documentada". Segons aquesta dimensió, la sèrie documental és tota aquella documentació, en qualsevol forma i suport, que genera una organització en el desenvolupament d'una funció o activitat pròpia. Aquesta dimensió és la que resulta de l'aplicació del principi arxivístic d'ordre originari.
- La **dimensió orgànica**, la qual prové de l'aplicació del principi arxivístic de procedència, i es basa en la relació que la sèrie manté amb l'òrgan que la genera. Aquesta dimensió intersecciona amb la funcional, de forma que ens permet definir la sèrie documental com aquella documentació, en qualsevol forma i suport, que genera un determinat òrgan d'una organització en el desenvolupament d'una funció o activitat pròpia.

Si col·loquem la dimensió funcional en l'eix d'abscisses, i la dimensió orgànica a l'eix d'ordenades, obtenim una matriu que ens permet identificar totes les sèries d'un organisme, i al mateix temps posar-les en relació tant per la dimensió funcional com per l'orgànica. Aquesta matriu és el que, en el marc d'aquesta tesi, i concretament en la guia de mapatge (**GMAP**), hem anomenat **mapa de sèries documentals** (Serra, 2010, 2013; Serra & Casellas, 2007).

Reconèixer la interseccionalitat de la sèrie documental, en la línia que proposa MoReq2010 en l'apartat "*Classification service*", permet configurar una visió de les agrupacions documentals més integrada en la gestió de l'organització, i posar en relació agrupacions que en una lectura merament jeràrquica no es podrien relacionar.

Figura 2-3: Perspectiva d'un tràmit només a partir de la jerarquia d'agrupacions documentals

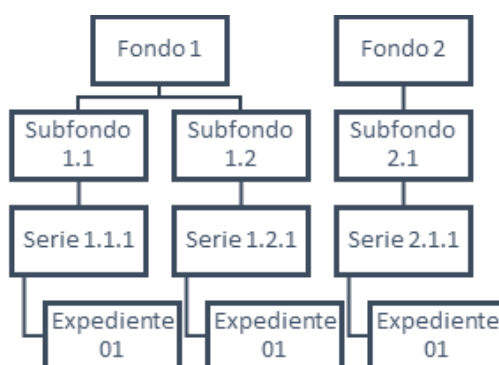
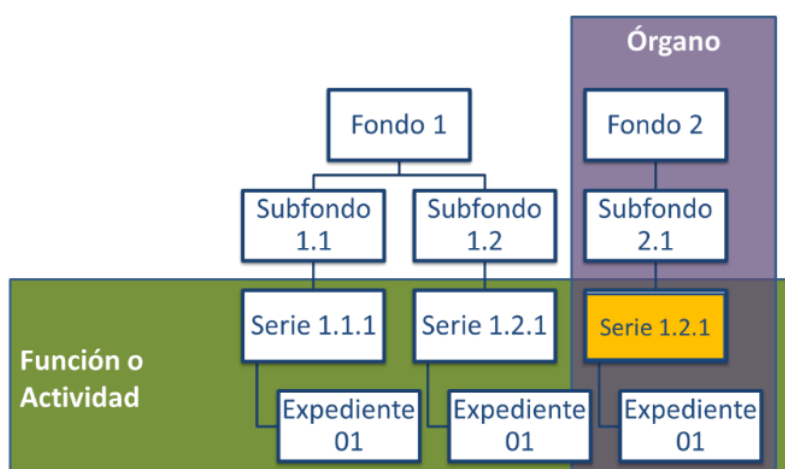


Figura 2-4: Perspectiva d'un tràmit a partir de la intersecció entre la funció i l'òrgan



La sèrie documental sorgeix doncs de la intersecció d'una taxonomia funcional (la variable transversal o horitzontal) amb l'estructura orgànica de l'organisme productor dels documents (la variable vertical) (Shepherd & Yeo, 2003; Yeo, 2009; Barbadillo, 2010) Aquesta suma dona com a resultat una matriu o mapa que posa en relació tots els elements o unitats d'informació que gestiona l'organització (el mapa de sèries documentals), i permet determinar, en funció d'aquestes relacions, el valor exacte de cada document dins del fons documental. Aquest mapa facilita la definició de calendaris de conservació, eliminació, transformació i accés als documents de cada sèrie documental, així com la definició de permisos i nivells de seguretat, juntament amb l'assignació de comportaments i fluxos de treball.⁹³

⁹³ Per poder identificar adequadament les sèries documentals, tota organització necessita doncs disposar de tres instruments: una relació de tots els agents de l'organització que poden actuar com a productors (el quadre de procedències), una relació de totes les funcions i activitats que pot portar a terme l'organització en l'exercici de les seves competències o interessos (el quadre de classificació), i una relació de totes les sèries documentals que en el moment actual l'organització està generant. En el cas d'estudi que s'ha desenvolupat en aquesta tesi, aquests

Des de principis del segle XXI s'han utilitzat tècniques de mapatge per a la identificació i normalització del substrat documental de les organitzacions, inspirades sovint en tècniques ja utilitzades en el desenvolupament d'aplicacions informàtiques (Keng & Xian, 2006; Rittgen, 2006; Stine et al., 2008) o en l'anàlisi forense digital (Ball, 2011). El precedent de referència és la metodologia *Infomap*, proposada per Cornelius Burk i Forest W. Horton a finals dels anys 80 (Burk & Horton, 1988). A nivell estatal, Carlota Bustelo i Elisa García-Morales introduïren el concepte de “mapa documental” en la consultoria de gestió documental l'any 1999 (Bustelo, 2012; García-Morales, 2014), el qual ha tingut seguiment en països de parla hispana (Zapata, 2002; Lloveras, 2014).

El mapatge de recursos d'informació s'ha desenvolupat de la mà de les ontologies i en forma de *knowledge graphs*, esdevenint una visualització cada cop més freqüent en els projectes de governança de la informació i, en alguns casos, de gestió documental (Barros & Sousa, 2019; Schmid et al., 2019; Smajevic et al., 2021; Smajevic & Bork, 2021; Chessa et al., 2022). Aquesta és la perspectiva que s'ha adoptat en el marc d'aquesta tesi.

2.3.1.4 Tipus de sèries documentals

L'arxivística clàssica ha elaborat diferents tipologies de sèrie documental, en funció de diferents criteris (Alberch et al., 2009). Analitzarem a continuació els tipus més habituals.

2.3.1.4.1 Segons el tipus d'expedient que genera

Definir el concepte d'expedient és un dels punts claus de l'arxivística, i un dels elements que d'una forma més clara diferencia l'arxivística de la resta de disciplines de la documentació pel que fa al tractament documental.

L'expedient es pot definir com la instanciació de la sèrie documental. Allò que la sèrie documental defineix (un context funcional, un cicle de vida, una composició documental, etc.) es materialitza en forma d'expedients (Shepherd & Yeo, 2003).

El concepte d'expedient ha rebut noms diferents. Tradicionalment la gestió per expedients ha estat vista com quelcom exclusiu del sector públic, i d'unes tradicions administratives concretes. Però l'àmbit privat ha utilitzat aquest concepte abastament, sota el concepte genèric de “*case management*” (De Man, 2009), i en el seu moment vinculat a l'ús d'eines de *workflow* i BPM (Van Der Aalst et al., 2005). El nom que ha

tres instruments s'han denominat mapa d'agents, mapa de processos i mapa de sèries documentals, i s'en detalla la construcció a la **GMAP** (apartat 3.2 del capítol 3).

rebut ha tingut relació tant amb la forma de tramitació (“*case handling*”) com amb la forma tradicional d’arxivar-ho (“*case file*” o “*case folder*”) (De Man, 2009).

El concepte d’expedient s’ha posat en relació de forma preferent amb la documentació d’activitats reglades, és a dir aquelles en les quals es coneix des de l’inici els passos que s’han de realitzar, la seva seqüència/ordenació i, en conseqüència, els documents que acabaran formant part de l’expedient (Alberch et al., 2009). Sheperd i Yeo anomenen aquestes activitats “*routine processes*”. Aquest coneixement previ prové del fet que existeix una regla, sovint documentada en forma de precepte legislatiu o procediment escrit, que determina aquesta seqüència d’accions i documents.

Però el concepte d’expedient pot abastar també la resta de l’espectre funcional d’una organització, que inclou totes aquelles tramitacions no reglades o de curs indeterminat, que corresponen a una activitat que es delimita en el temps (comença i acaba) però que no es desenvolupa segons un esquema predefinit (Alberch et al., 2009). Són les que Sheperd i Yeo anomenen “*creative processes*” (Shepherd & Yeo, 2003). Això fa impossible predir la composició de l’expedient, almenys de forma exacta i ordenada. Aquest segon tipus d’expedient s’ha anomenat sovint d’altres formes per a diferenciar-lo de l’expedient reglat, com per exemple *dossier*, *case file*, etc., i s’ha identificat amb una simple carpeta, creada per motius circumstancials per agrupar temporalment documents.

Tant Foscarini com Sheperd i Yeo consideren que és molt més senzill identificar els processos reglats, que tenen naturalesa clarament repetitiva, que no pas els no reglats, la identificació dels quals incorpora elements més subjectius i imprecisos (Shepherd & Yeo, 2003; Foscarini, 2012). Foscarini considera que hi ha parts de l’activitat humana que no poden ser reduïdes a processos repetitius (Foscarini, 2012).

Hi ha un tercer tipus d’activitats, que més difícilment es poden agrupar en expedients. Es tracta d’aquelles activitats que per la seva durada son tan curtes que amb prou feines donen origen a més d’un document. I a vegades ni tan sols això, ja que es redueixen a anotacions en registres o bases de dades. Aquestes actuacions només es poden agrupar en unitats superiors a efectes pràctics, ja que des d’un punt de vista estrictament funcional corresponen a activitats que es delimiten per sí soles en el temps, i no tenen relació causa-efecte amb d’altres actuacions. Aquestes agrupacions es poden fer en expedients compendi o temàtics, corresponents a períodes cronològics en el desenvolupament de determinades funcions (per exemple totes les invitacions refusades durant el mes de gener), o bé en els propis registres o bases de dades, que reben aleshores el nom de documents recapitulatius (Alberch et al., 2009).

Aquest tercer tipus d’expedient és el punt per on sovint es trenca el criteri funcional en l’agrupació de documents, i es torna a plantejaments més propis d’arxius històrics, com l’agrupació dels documents per temàtica, objecte de l’activitat o fins i tot forma externa dels documents (tipus documental) (Gallarreta et al., 2012). Aquests expedients, anomenats *subject files*, en funció de la llunyania amb el criteri funcional podran seguir

conformant sèries documentals, en aquest cas anomenades factícies (mentre conservin encara un vincle funcional, com en el cas de les tipologies documentals o els objectes) o bé passaran a conformar col·leccions (en el cas de les agrupacions purament temàtiques) (Shepherd & Yeo, 2003).

Aquesta tercer tipus d'expedient porta a plantejar-se quina és la utilitat real del concepte d'expedient, especialment en l'entorn digital. De fet, aquesta utilitat es pot valorar des de dos punts de vista:

- **L'expedient com a entitat documental amb significació pròpia.** Quan parlem d'expedient, no parlem d'una entelèquia o construcció fictícia, sinó d'una entitat que en l'àmbit arxivístic té una significació pròpia (Yeo, 2012). Això vol dir que té uns atributs propis que poden ser objecte de descripció, i que es correspon amb una realitat que existeix en totes les organitzacions: l'activitat o tramitació d'un assumpte (Van Der Aalst et al., 2005). Probablement pot ser objecte de debat que tots els documents d'una organització s'agrupin en expedients, però no que existeixi de forma objectiva l'entitat expedient, especialment quan es parla d'expedients reglats i no reglats (tramitacions i activitats). I quan els documents deixen d'existir com objectes, l'expedient pot seguir actuant com a evidència d'una activitat, d'acord amb la polisèmia del terme "*record*" en l'entorn anglosaxó (Yeo, 2011).
- **L'expedient com a recurs de gestió.** Amb independència de la seva significació documental, el cert és que el concepte d'expedient és un recurs cabdal per facilitar la gestió de la documentació administrativa a les organitzacions i reduir-ne la seva complexitat (Alberch et al., 2009; Beach & Oates, 2014).

D'acord amb l'exposat, podem convenir l'existència de tres tipus d'expedients, amb les característiques que s'indiquen a continuació, les quals s'extrapolen a les sèries documentals que els contenen:

Taula 2-4: Característiques dels tipus d'expedient

Tipus	Procediment	Composició	Ordenació interna	Límits temporals	Caràcter de la sèrie docum. ⁹⁴
Reglat	Regulat	Coneguda o previsible	Previsible	Coneguts i originals	Preferentment primari
No reglat tècnic	No regulat	Coneguda o previsible	Imprevisible	Coneguts i originals	Preferentment primari
No reglat temàtic	Inexistent	Desconeguda o imprevisible	Imprevisible	Convencionals	Preferentment secundari

⁹⁴ Per a la diferenciació entre sèrie primària i secundària vegeu l'apartat 2.3.1.4.2.3 d'aquest capítol.

Des de la comprensió de l'expedient com un recurs per a la gestió, el coneixement del tipus d'expedient, i per tant del tipus de sèrie documental, ens permet determinar l'entorn preferent d'utilització, aquell en el qual la seva aparició serà més habitual (Serra & Casellas, 2007):

Taula 2-5: Entorns d'aplicació dels tipus d'expedient

Tipus	Propi de...	Formes documentals on s'aplica
Reglat	Unitats de tramitació	Conjunts de documents i/o matrius de dades transaccionals
No reglat tècnic	Unitats tècniques	Conjunts de documents
No reglat temàtic	Unitats directives o funcions no pròpies	Conjunts de documents, bases de dades i aplicacions

En funció dels tipus d'expedient que hem descrit podem doncs classificar les sèries documentals. Les sèries que responen a funcions o activitats, i amb uns límits coneguts (reglades i no reglades tècniques) es consideren sèries “naturals”, perquè són el resultat directe d'una activitat, la qual documenten de forma precisa. També les podem denominar sèries transaccionals, una denominació que utilitzarem en aquesta tesi. Per oposició a aquestes sèries, parlarem de sèrie recapitulativa per definir aquella sèrie que recopila fraccions de sèries transaccionals en una forma més compacta, i habitualment amb caràcter registral, és a dir d'evidència immodificable. I reconeixem un tercer tipus de sèrie, aquella que no respon a un criteri de classificació funcional, o coherent amb el criteri majoritari en aquell sistema de gestió documental, i que s'ha creat per pura practicitat, però sense una significació documental específica, de forma “artificial”. Són les sèries anomenades factícies, també denominades col·leccions en l'àmbit dels arxius històrics, i pateixen d'una notable indefinició en la literatura arxivística (Delgado, 2007). Tot i així, alguns autors neguen la distinció entre agrupació “natural” i “artificial”, atès que el propi fet de reflectir una activitat amb una determinada seqüència documental, predeterminada per l'arxiver, també té un component artificial (Yeo, 2012).⁹⁵

⁹⁵ L'arxivística ha utilitzat el criteri funcional com una forma de diferenciar les agrupacions documentals arxivístiques, que s'han anomenat “naturals”, de les agrupacions realitzades en biblioteques, basades en col·leccions, que per oposició els arxivers han anomenat “artificials” (Yeo, 2012).

2.3.1.4.2 Segons les relacions que mantenen entre elles

Un dels debats que s'ha generat al voltant de les classificacions arxivístiques ha estat el derivat d'intentar ubicar la sèrie documental dins una estructura de classificació habitualment jeràrquica o taxonòmica (Foscarini, 2012). En alguns casos, la sèrie s'ha ubicat en un mateix nivell de la jerarquia, com és el cas del Sistema General de Gestió de la Documentació Administrativa de la Generalitat de Catalunya,⁹⁶ i en d'altres s'ha considerat present en qualsevol nivell de l'estructura de classificació, obrint la porta a parlar de sèries i subsèries (Rodríguez, 2012). En tot cas, la inclusió de la sèrie en algun nivell d'una taxonomia implica l'atribució a la sèrie d'una relació de comprensió, que és la que ha contemplat RiC per a l'element *Record*.

Però una lectura de la sèrie des de la perspectiva del procés que la genera permet contemplar altres tipus de relacions entre les sèries, tenint present la riquesa de relacions que els processos poden mantenir entre ells (Rosing, 2014), i que alguns autors ja han destacat en relació amb la classificació funcional (Aparecida et al., 2012; Sundberg, 2013; Niu, 2014). Sorprenentment, les relacions no jeràrquiques entre sèries documentals són un dels aspectes que menys interès ha despertat en la literatura arxivística, i sobre els quals la disciplina pràcticament no ha creat doctrina (García Ruipérez, 2007; Rodríguez, 2012; Yeo, 2009, 2012). Mariano García Ruipérez es feia ressò d'aquest buit analitzant les pràctiques en l'estat espanyol, i posant de relleu la manca de consens en la conceptualització de les relacions entre sèries (García Ruipérez, 2007), i Mokhtar i Yusof ho assenyalen com una de les mancances de la classificació funcional (Mokhtar & Yusof, 2015).

En el context d'aquesta tesi hem contemplat cinc tipus de relacions entre sèries, de concepció pròpia, que enumerem a continuació:⁹⁷

2.3.1.4.2.1 Relació estructural (funcional)

Un primer tipus de relació és la que mantenen les sèries generades per tots els òrgans que participen en una mateixa funció o macroprocés. Correspon al desplegament jeràrquic de la taxonomia funcional, i és sovint la única relació que recullen molts quadres de classificació que s'utilitzen actualment.

Però es poden identificar quatre tipus més de relació entre sèries documentals, en funció del procés o activitat que genera les sèries, del moment en què es formen, de l'agent

⁹⁶ Vegeu <https://cultura.gencat.cat/ca/temes/arxiu-i-gestio-documental/gestio-documental/qdgcencat/aresta/protocol/instruments/qdcfgencat/index.html>.

⁹⁷ La conceptualització d'aquestes relacions és deutora de converses mantingudes amb Lluís-Esteve Casellas, de l'Ajuntament de Girona, i contrastades amb d'altres professionals del col·lectiu arxiver català. Malauradament no s'han localitzat publicacions que les formulin en aquests termes.

que les produeix, i de l'objecte de l'expedient. Detallem a continuació cadascun d'aquests tipus de relació.

2.3.1.4.2.2 Relació de continuïtat (antecedent/conseqüent)

Aquesta relació correspon a la relació que mantenen les sèries que procedeixen d'activitats que s'encadenen entre elles per una relació de causa-efecte. La relació de continuïtat és doncs aquella que manté una sèrie amb les sèries que provenen d'activitats immediatament anteriors (antecedents) o posteriors (conseqüents), amb les quals manté una dependència que en determina el contingut.

2.3.1.4.2.3 Relació de derivació (primària/secundària)

Es defineix la relació de derivació com aquella que mantenen les sèries que són producte del mateix procés o activitat. En aquest cas, la diferència ve donada per l'agent productor, i pel nivell de competència sobre el procés que tingui cada agent.

Aquesta relació es dóna en processos o activitats que són competència exclusiva d'un agent, i on hi participen d'altres agents amb funcions d'assessorament, suport o control. En aquest cas, la sèrie documental que genera l'agent que té la competència s'anomena **sèrie primària**, i les dels agents no directament competents, **sèries secundàries**.

2.3.1.4.2.4 Relació de dependència

No és habitual que la competència sobre un procés estigui compartida entre diferents agents. Si es dóna aquest cas, totes les sèries d'agents competents tindran la consideració de sèries primàries, i el tipus de relació que establiran serà una relació de dependència. En aquest cas les sèries relacionades contindran documentació diferent del mateix procés o activitat, ja que cadascun dels agents que hi participen exerceix, en l'àmbit en què participa, una competència plena.

Hi ha casos però en què el desenvolupament del procés pot implicar que es portin a terme, per part de diferents agents, accions o transaccions relacionats i complementaris, però de característiques diferents. En aquest cas, parlarem igualment de dues sèries relacionades, però no per la pertinença a un mateix procés, sinó per una relació de complementarietat entre dues activitats. Aquesta és la relació que mantenen entre elles sèries que són producte de diferents processos o activitats, però el valor de les quals depèn de la sèrie relacionada. Aquesta relació és especialment rellevant pel tema que tractem en l'apartat 2.3.3 d'aquest capítol, atès que introdueix el concepte de

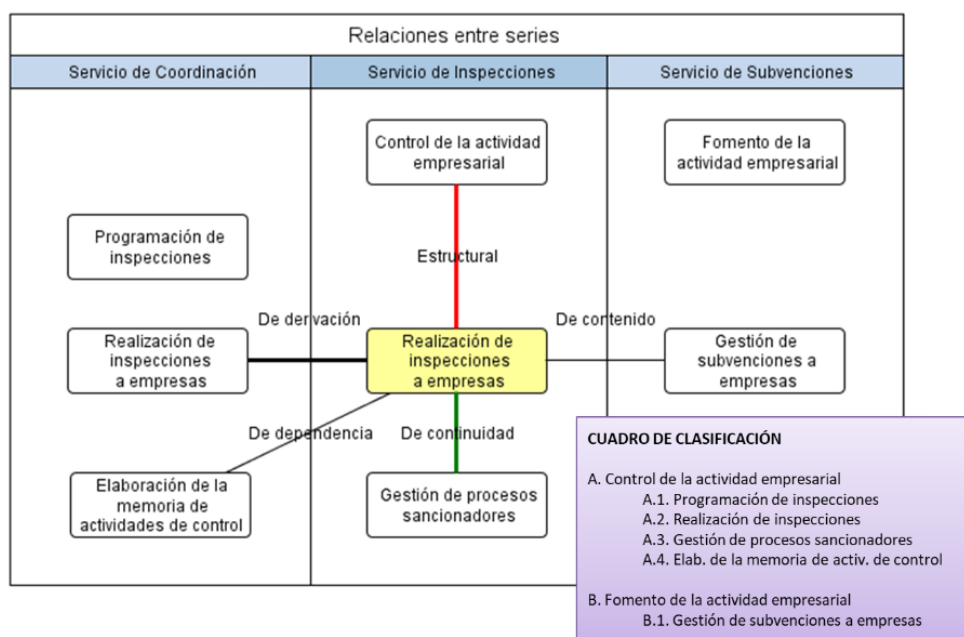
transversalitat o recapitulació, trencant l'estructura jeràrquica típica de moltes classificacions funcionals.

2.3.1.4.2.5 Relació de contingut

És la relació que s'estableix entre sèries que no mantenen cap vincle derivat del seu procés o activitat, però que fan referència al mateix objecte o destinatari. En aquest cas, l'únic que tenen en comú les sèries és la seva temàtica o objecte. Aquesta relació es manifesta habitualment en la construcció del criteri de nominalització dels expedients de la sèrie.⁹⁸

En el següent diagrama es resumeixen aquestes relacions:

Figura 2-5: Representació de les possibles relacions entre sèries documentals



2.3.2 El requisit d'existència de la sèrie documental: el context funcional

El principal focus de debat amb la sèrie documental sempre ha estat el vincle amb el context que la genera, és a dir els processos o les activitats, el que ha determinat en

⁹⁸ Per una definició del concepte de criteri de nominalització, vegeu la descripció de l'atribut ENT007-003 Criteri de nominalització, a l'apartat 4.6.1.7.4 del capítol 4.

conseqüència la construcció dels sistemes de classificació (Mokhtar & Yusof, 2015; Mokhtar & Zawiyah, 2015; Packalén, 2015).

El quadre de classificació ha estat un instrument controvertit (Ifould & Joseph, 2016). L'arxivística clàssica, i més vinculada als arxius històrics, havia considerat sempre el quadre de classificació com un instrument de descripció, on es descrivien entitats documentals o amb significació documental (Moyano, 2013). El fet que siguin entitats documentals implica que tenen una existència física, i que, en conseqüència, la seva descripció no pot fer-se a priori; la identificació de sèries només pot provenir de l'existència prèvia d'aquestes sèries. El quadre de classificació podia, en nivells superiors, incloure informació dels productors de les sèries, és a dir del fons i subfons en els quals s'incardinaven les sèries. Per aquest motiu els quadres de classificació descriptius rebien sovint el nom de quadres de classificació orgànico-funcionals (Alberch et al., 2009).

Però la gestió documental va trencar la idea del quadre de classificació com a instrument de descripció (Roberge, 1992). El model que proposava el *records management* no era el d'un instrument de descripció, sinó el d'un instrument per a la descripció. És a dir, un instrument per normalitzar la descripció d'entitats documentals, cosa la qual implica que ha de tenir una factura apriorística, prèvia a l'existència física de les entitats documentals. El criteri en base al qual es plantejava normalitzar la denominació de les entitats documentals, especialment al nivell de sèrie documental, era el corresponent a l'ordre originari de creació de l'entitat, és a dir al desenvolupament de la funció i activitat que donava origen a aquells documents. Per aquest motiu, el quadre es va anomenar quadre de classificació funcional.

Quan la classificació funcional es va anar estenent, alguns autors van intentar adaptar el concepte de sèrie a una estructura funcional, utilitzant per això el ja existent nivell de subsèrie combinat amb un nivell superior de sèrie documental "genèrica" (Rodríguez, 2012). També diferenciant entre sèries comunes, i sèries específiques, de la mateixa manera que ho feien els quadres de classificació funcionals de procedència canadenca (Roberge, 1992; *Guia de classificació de documents del sistema general de gestió de la documentació administrativa*, 2015).

Amb el temps les dues visions del quadre de classificació s'han anat diferenciant. El model MoReq (MoReq2010, 2010-2011) diferencia entre la necessitat per a les organitzacions de disposar d'un *business classification scheme* (o quadre de classificació de negoci, de caràcter funcional) i d'un *records classification scheme* (o quadre de classificació d'entitats documentals, amb caràcter d'instrument de descripció). I de fet la implementació clàssica dels sistemes de gestió documental en eines EDMS s'ha basat durant molt de temps en la creació d'una taxonomia de classificació aplicable per metadades (el *business classification scheme*), i un desplegament d'aquesta

taxonomia en una estructura físico-lògica⁹⁹ de carpetes, l'anomenat "*file plan*" (Jones, 2008; Ifould & Joseph, 2016).

L'arxivística anterior a l'aparició de la gestió documental no es plantejava una intervenció activa en la creació dels documents, i molt menys determinar-ne aspectes formals o d'autenticitat (Oladejo & Hadžidedić, 2021). El patró de treball consistia en analitzar les característiques dels documents rebuts a l'arxiu, i a partir d'aquests derivar-ne el criteri de formació de la sèrie documental, als efectes bàsicament de descripció.¹⁰⁰ Els models més inspirats en l'arxivística clàssica (recollits en la norma ISAD-G) han tendit a identificar la sèrie documental a partir del tipus documental predominant, és a dir a partir de les característiques formals, estructurals i de significació dels documents (García Ruipérez, 2007; Bouth & Moreira, 2022).

Per aquest motiu, sovint s'ha utilitzat el terme "tipus documental" enlloc de sèrie documental per a definir el criteri d'agrupació dels documents en expedients.

2.3.2.1 La classificació per funcions i activitats

Des del punt de vista de l'arxivística tradicional, la identificació de funcions o activitats, entesa com a mètode de classificació i ordenació, s'ha associat a un dels dos principis axials de l'arxivística, el principi de l'ordre originari, formulat per M. Lehmann el 1881 (Niu, 2015). El principi ha admès durant molt temps una lectura bàsicament orgànica, entenent l'estructura de producció dels documents com una estructura d'unitats administratives o operatives (Henttonen & Kettunen, 2011). Però a partir de l'eclosió del *records management* aquest principi va prendre una dimensió funcional, entenent l'estructura productora com el conjunt d'agents i accions que porten a la construcció d'una determinada agrupació documental (Mokhtar & Yusof, 2015).

La classificació per funcions i activitats es justifica per la pròpia naturalesa dels documents com a productes de les accions (Schellenberg, 1956; Foscarini, 2012). L'èmfasi en la importància del context funcional ha determinat la classificació en l'arxivística contemporània, i les tècniques d'anàlisi de funcions, activitats i processos han esdevingut la base de l'elaboració de les classificacions arxivístiques (Sabourin, 2001; Orr, 2005; Kettunen & Henttonen, 2010; Foscarini, 2012).

Però malgrat la importància que l'arxivística li ha donat a la identificació de funcions i activitats, ha desenvolupat poques metodologies solvents per a portar a terme aquesta identificació, o per a resoldre problemes conceptuals com el de les sèries documentals

⁹⁹ Jones feia referència al concepte de "*virtual folder*", una recreació per a l'usuari d'una carpeta d'expedient utilitzant metadades i vistes dins d'un EDMS (Jones, 2008).

¹⁰⁰ Per aquest motiu, durant molt temps en l'àmbit dels arxius històrics el concepte de quadre de classificació s'ha identificat amb una forma d'instrument de descripció, i no amb un instrument per a l'aplicació proactiva de cicles de vida sobre els documents.

transversals (Foscarini, 2006; Mokhtar & Yusof, 2015). Sovint els conceptes de funció o activitat (“*function terms*”) i els de temàtica o *topic* (“*subject terms*”) es confonen en les organitzacions, i en el procés d’identificar els nivells més genèrics (“*abstract functions*”) s’acaba identificant la temàtica amb la funció (Foscarini, 2012). Segons Fiorella Foscarini, la manca de sistemàtica en la classificació arxivística, el fet que no s’hagin desenvolupat metodologies potents, és el resultat de que sigui el productor el responsable d’organitzar els documents (Foscarini, 2006).

Tot i considerar-s’hi oposada, la classificació funcional no s’ha pogut deslliurar de la classificació orgànica, sobretot mentre l’estructura de les organitzacions ha estat fortament jerarquitzada, i això ha propiciat el caràcter monojeràrquic de la majoria de classificacions funcionals arxivístiques (Foscarini, 2012).

Als anys 90 l’arxiver canadenc Michel Roberge, arran del primer Màster d’Arxivística organitzat per l’aleshores Associació d’Arxivers de Catalunya, va portar a Catalunya el concepte de classificació funcional. Aquest concepte fou innovador per als arxivers catalans especialment en dos aspectes:

1. El primer era l’homogeneïtat en el criteri de classificació. Tota la taxonomia emprada per classificar era íntegrament funcional, a diferència dels anteriors instruments de descripció arxivístics, on les estructures de classificació podien combinar criteris diferents, i aquests corresponien habitualment, en els nivells superiors, a un criteri orgànic.
2. La sèrie documental no era un nivell específic en el quadre de classificació. A imitació dels sistemes de classificació bibliotecaris, el quadre de classificació era un instrument independent dels instruments de descripció, i qualsevol entrada del quadre, a qualsevol nivell de la jerarquia, podria “obrir sèrie”.

Aquesta idea és la que es va aplicar durant molt de temps a Catalunya (Alberch et al., 2009). També és la idea que va recollir el DLM Forum de la Unió Europea (Vieira et al., 2012). MoReq2010 obvia en el seu esquema el concepte de sèrie, i defineix l’entrada d’un quadre de classificació com “*Classe (class): fragment d’una jerarquia representat per una línia que va des d’un punt de la jerarquia del quadre de classificació fins a tots els expedients inferiors a aquest punt. Això pot correspondre, en terminologia clàssica, a “classe primària”, “grup” o “sèrie” (o subclasse, subgrup subsèrie, etc.) en qualsevol nivell del quadre de classificació*”.

Nombroses definicions reforcen aquest vincle funcional del concepte de sèrie documental. Elisabeth Shepherd i Geoffrey Yeo fan una definició de sèrie compacta però molt efectiva: Una sèrie comprèn “*the records of all the activities that are instances of a single process*” (Shepherd & Yeo, 2003).

El criteri funcional ha esdevingut, per tant, el criteri preferent, i gairebé únic, per a la classificació de documents en la gestió documental (Milne, 2007). La publicació de la norma ISO 15489, a partir de la norma australiana AS4390 (Healy, 2001), va consagrar

a nivell d'estàndard aquest requisit, i a partir d'aquí ha estat present en la doctrina arxivística de forma general. Tots els models de descripció incorporen aquesta visió, i fixen la relació de funcions i activitats de l'organització com la gènesi de qualsevol agrupació documental.¹⁰¹

Un pas molt important per a la sistematització d'aquest origen funcional de la sèrie documental es va donar amb la norma ISO 26122, hereva del mètode australià DIRKS-B. La norma ISO 15489 fixava una estructura funcional a tres nivells que ha quedat com a pauta de referència: la que diferencia entre funció, activitat i transacció. Aquesta divisió en tres nivells (funció, activitat i transacció), el que s'ha anomenat el model F-A-T (Foscarini, 2012), ja fou proposat per Schellenberg (Schellenberg, 1956, 1965), i ha marcat la pauta fins a dia d'avui de la classificació arxivística (Shepherd & Yeo, 2003). Però no fou fins la publicació de la norma ISO 26122 que es van fer dues aportacions molt rellevants:

1. Explicar com es poden identificar aquests nivells, quan cal fer-ho, i a quin tipus d'entitat corresponen realment.
2. Incorporar un concepte que, fora de l'àmbit arxivístic, és el concepte de referència per a la descripció de l'activitat de les organitzacions: el concepte de procés.

2.3.2.2 El concepte de procés i la identificació de sèries documentals

El concepte de procés no és un concepte propi de la teoria arxivística, ni és habitual en la seva bibliografia (Sundberg, 2013). De fet, el concepte de procés és l'eix de tres disciplines professionals directament relacionades amb la gestió de la informació: l'anàlisi i millora de processos, la gestió documental i l'enginyeria informàtica. Un element que va unir fortament aquestes tres disciplines fou la popularització cap a mitjan dels anys 90 de les solucions empresarials de *Workflow*, posteriorment denominades BPMS (*Business Process Management System*) (De Man, 2009).¹⁰² Les eines de BPM es recolzaven en eines EDMS (popularment anomenats gestors documentals), on desaven els documents i les metadades dels flux en execució, i sovint es distribuïen com a parts del mateix paquet o solució (Van Der Aalst et al., 2003). Molts arxivers i *records managers* entraren en contacte amb l'enginyeria de processos a través de l'ús de les eines de BPM integrades amb els gestors documentals.

És rellevant fer un repàs de les normes de descripció arxivística per observar com recullen de diferents formes el concepte de procés.

¹⁰¹ Vegeu el model de dades per a totes les correspondències de les normes de descripció amb els conceptes funcionals.

¹⁰² La *Workflow Management Coalition* es fa fundar l'any 1993. Vegeu <https://wfmc.org/>.

Dins el conjunt de normes del Consell Internacional d'Arxius anterior al *Records in Context*, trobem una norma dedicada monogràficament a la identificació i classificació de funcions i activitats, la norma ISDF: *International Standard for Describing Functions* (ISDF, 2007). Aquesta norma identifica tres nivells, que s'ordenen per nivell de concreció del més general al més concret: funció, activitat i transacció, sense contemplar entre els nivells el de procés com a tal.

En la darrera edició de la segona part de la norma ISO 23081 (ISO 23081-2:2021, 2021) es defineixen els conceptes a descriure dins un sistema de gestió documental, i en ells s'inclouen com a entitat els processos de l'empresa (*business processes*). Un subconjunt d'aquests processos, que la norma considera de forma destacada, però inclosos en els processos de l'empresa, són els processos de gestió documental (*records management business processes*). La norma que la va inspirar, l'*Australian Government Recordkeeping Metadata Standard* (AGRkMS). Versió 2.2 contemplava el concepte que comentem, per nivells similars a ISDF, però el desgranava en quatre nivells: *Ambient function*, *Business function*, *Business activity*, *Business transaction*.

En l'àmbit estatal, el model e-EMGDE inclou una entitat, anomenada "*Actividades*", que es desenvolupa en quatre nivells de progressiva concreció: "*Función marco*", "*Función*", "*Actividad*" i "*Acción*". Però aquests elements no es relacionen només jeràrquicament. Anticipant el caràcter ontològic de la línia de treball de RiC, preveu que les relacions entre aquests quatre nivells siguin de grau N a N, és a dir, que es puguin desenvolupar en forma de xarxa.

Aquest caràcter ontològic és el que ja va incorporar des del seus inicis l'altra línia de treball espanyola, la que va portar a terme el Ministeri de Cultura elaborant les *Normes Espanyoles de Descripció Arxivística* (NEDA). Membres de la comissió redactora de la NEDA col·laboraven també amb el Consell Internacional d'Arxius en l'elaboració del *Records in Context*, per la qual cosa ambdues normes tenen punts en comú. Però en el cas concret de les funcions i els processos, mentre que la NEDA inclou una entitat "*Función*" que es desplega en quatre nivells: "*Función*", "*División de función*", "*Actividad o proceso*" i "*Acción*", *Records in Context* contempla només una entitat, que anomena "*Event*", que es desenvolupa en un segon nivell corresponent a "*Activity*".

L'elaboració de classificacions funcionals ha posat en contacte els arxivers amb les tècniques de modelat de processos (Mokhtar & Yusof, 2016), i en conseqüència amb les indeterminacions d'aquestes tècniques. Segons Geoffrey Yeo la indefinició del concepte de procés devalua el caràcter "exacte" o "natural" de la classificació funcional. Citant al consultor Fred Nickols, Yeo afirma que "*analyses of processes are [...] adrift in a sea of undefined terms, unclear boundaries, different perceptions and experiences*" (Yeo, 2012). De fet, l'anàlisi de processos posa de relleu la dificultat de mantenir el concepte de "naturalitat" de les agrupacions documentals arxivístiques, sense que això impliqui menystenir la seva utilitat com a recurs de gestió des d'un punt de vista

purament pragmàtic (Bak, 2012). Tant els límits dels processos com les relacions entre els processos és quelcom que s'acaba determinant de forma convencional (Yeo, 2012).

El concepte de procés està formalitzat des dels anys 70 (Serifi et al., 2013), i es defineix com el conjunt d'activitats d'un o diversos agents adreçades a aconseguir un objectiu comú, generades per una entrada (*input*) d'un proveïdor, i orientades a la producció d'un producte o servei (*output*) cap a un client (intern o extern). Els factors que condicionen el procés en aquest model són (Rosing, 2014):

- El factor subjectiu: els agents que intervenen en el procés.
- El factor objectiu: aquella cosa que el procés produeix o transforma.

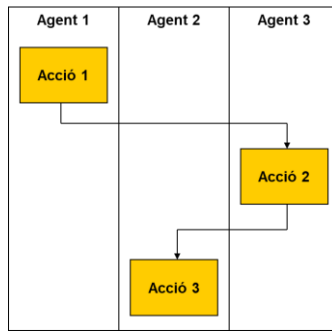
El fet que el procés tingui una entrada i una sortida és el que permet encadenar processos, i establir relacions de continuïtat entre els processos, quan la sortida d'un procés es converteix en l'entrada de l'altre. Això acaba fixant un conjunt de relacions que podríem denominar estructurals, pel fet que estructurin funcionalment tota l'organització. Però al mateix temps esdevé el principal factor de complexitat en el moment d'identificar processos, atès que els límits del procés (*l'input* i *l'output*) són objectes en transformació, aparentment limitada en el temps, però que en determinats casos es perllonga de forma continuada durant tota l'activitat de l'organització.¹⁰³ El tall entre un procés i un altre esdevé en aquests casos una decisió convencional, i per tant exposada a subjectivitat (Yeo, 2012; García-Morales, 2014). Per evitar aquesta indefinició en la identificació de processos s'ha fet diversos esforços per predeterminar l'estructura de processos d'una organització, la més coneguda de les quals diferencia entre processos estratègics, processos clau o operatius, i processos de suport (Rosing, 2014).

Dins els processos podem identificar múltiples models de procediment,¹⁰⁴ però els més habituals són dos (Van Der Aalst et al., 2005). El primer és el **procediment seqüencial**, on les tasques s'executen seguint una seqüència predefinida, el flux de treball o *workflow* (Van Der Aalst et al., 2003). El procediment seqüencial inclou habitualment bifurcacions en funció del compliment de determinades condicions (*gateways* inclusivament, exclusivament, paral·lels, etc.), i involucra en aquesta seqüència diversos agents. La determinació de l'agent principal en aquest disseny ve determinada habitualment per la competència, reconeguda organitzacionalment, de l'agent sobre el procés.

¹⁰³ Si agafem per exemple els processos d'escolarització, observarem com el procés de preinscripció finalitza amb un alumne assignat a un centre, que posteriorment esdevé alumne matriculat en aquell centre, i posteriorment alumne avaluat, i posteriorment alumne graduat.

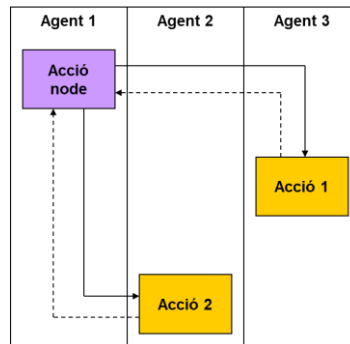
¹⁰⁴ Entenent procediment com l'ordre en què s'executen les accions d'un procés.

Figura 2-6: Esquema del model de procediment seqüencial



El segon és el **procediment radial**, on no existeix cap seqüència predefinida en l'execució de les accions, i és la voluntat de l'agent principal la que actua de disparador de cada branca d'execució, la qual cosa no exclou l'existència de regles d'obligatorietat o dependència en l'execució de les diferents accions. En aquest cas l'agent principal ve identificat pel propi flux del procediment. Aquest segon model, denominat també *case handling model* (Van Der Aalst et al., 2005), equival a una gestió per expedients tradicional, on l'expedient té un responsable que condueix el llençament de les accions en funció de cada context. Aalst, Weske i Grünbauer definien tres possibles rols en aquest sentit: el tramitador (*execute role*), i dos nivells de supervisor, el que pot modificar regles (*redo role*) i el que pot obviar tasques (*skip role*) (Van Der Aalst et al., 2005).

Figura 2-7: Esquema del model de procediment radial

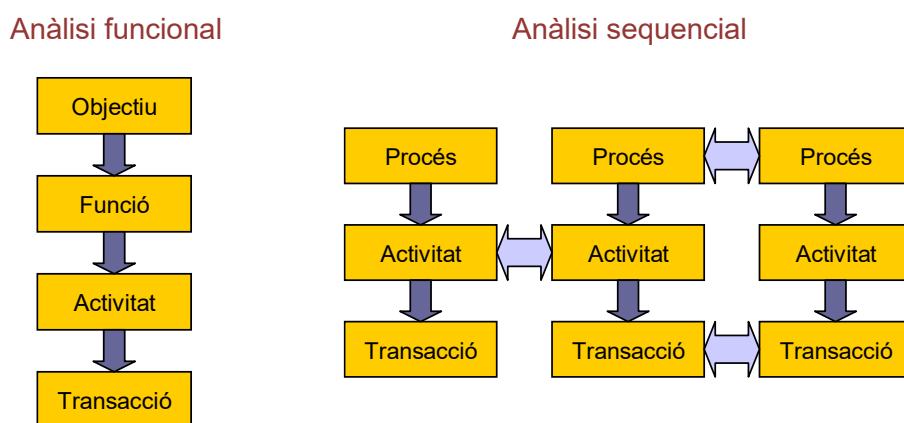


2.3.2.3 La jerarquitització dels processos i l'impacte sobre la classificació funcional

El més habitual en un mapa de processos és estructurar els processos en diferents nivells, amb una estructura taxonòmica, fet que ha inspirat la classificació funcional (Mokhtar & Zawiyah, 2015; Mokhtar & Yusof, 2016).

La norma ISO 26122 proposa dos tipus d'anàlisi per a la construcció de la classificació funcional. El primer és l'anàlisi funcional o jeràrquic, que consisteix a descompondre en parts lògiques les funcions d'alt nivell de l'organització, a partir de la identificació d'un marc general de funcions de l'organització que el mètode DIRKS anomena *big picture* (DIRKS-B, 2001). El segon tipus és l'anàlisi seqüencial o de processos, el qual parteix de l'anàlisi en detall de com l'organització desenvolupa les seves activitats, per posar progressivament en relació aquestes activitats fins a construir el mapa de processos o funcional. El model DIRKS (DIRKS-B, 2001) incorpora elements de tots dos tipus d'anàlisi.

Figura 2-8: Esquema dels dos models d'anàlisi de la norma ISO 26122



A diferència de l'enfocament funcional, que és bàsicament vertical, l'enfocament a processos és relacional. De fet, en un enfocament a processos pur la transversalitat és l'element definidor de l'estructura (Aparecida et al., 2012). Aquest és l'enfocament que hem adoptat en el model que s'ha aplicat al Consorci d'Educació de Barcelona, i el que millor respon al model interseccional de sèrie documental propi del sistema australià de sèries (Cunningham et al., 2013; Hurley, 2008).

La bibliografia arxivística, malgrat haver fet referència als processos, probablement no els ha integrat com un element de transformació dels seus sistemes de classificació, restant bàsicament en el nivell funcional de la ISO 26122 (Alonso & Lloveras, 2007; Escobar, 2009; García-Morales, 2014; Sierra, 2012). Fiorella Foscarini arriba a afirmar que els arxivers no acaben de tenir clar què s'entén per funció, activitat i transacció, i per aquest motiu moltes classificacions arxivístiques manquen de coherència, cosa la qual ha acabat restant operativitat a la classificació funcional (Foscarini, 2012).

2.3.3 La crisi del model funcional

Des de fa una dècada es multipliquen les visions crítiques amb els sistemes de classificació arxivístics, especialment amb el predominant sistema funcional. Aquestes crítiques venen motivades per dos factors:

- En primer lloc, l'evolució tecnològica ha obert consecutivament opcions per estructurar, gestionar i recuperar la informació que superen àmpliament en prestacions la tradicional visió monojeràrquica dels quadres de classificació arxivístics.
- I en segon lloc, el propi criteri funcional, que està en la base de la metodologia de la gestió documental, que es considera la clau per donar context als documents i sentit a les agrupacions, i que és establert com preceptiu per la norma ISO 15489, està començant a ser qüestionat.

Analitzem doncs cadascun dels dos plantejaments.

2.3.3.1 Crítica de la classificació jeràrquica i única

Les funcionalitats dels gestors documentals i l'evolució de la gestió documental electrònica va portar, fa aproximadament una dècada, a un replantejament de la classificació arxivística, i en conseqüència del concepte d'agrupació documental (Oladejo & Hadžidedić, 2021). El debat girava al voltant de la regla, interioritzada durant molt temps per la comunitat arxivística, i consagrada en la major part de models de descripció, que establia que els documents només poden ser classificats d'acord amb un sol criteri, i per tant només poden tenir un únic codi de classificació. Aquesta regla implica doncs que els documents només poden pertànyer a una única agrupació documental.

Els gestors documentals electrònics oferien dues noves possibilitats:

1. La primera era la possibilitat d'accedir a un mateix document des de diferents punts. La separació interna dels EDMS entre un sistema de fitxers i una base de dades que contenia les metadades i els apuntadors als fitxers permetia gestionar aquesta opció. El document, per tant, podia aparèixer i ser utilitzat, a nivell operatiu, en contextos diferents, sense deixar de ser físicament el mateix fitxer, ni perdre els mitjans d'autenticació que en garantien l'autenticitat. Apareixia doncs el concepte de "document compartit".
2. I la segona era la possibilitat d'associar a cada document un nombre gairebé infinit de metadades, que obria la possibilitat, amb les metadades categòriques, a classificar un mateix document segons diversos criteris.

Aquesta nova situació va portar a alguns autors, sobretot de l'escola canadenca, a proposar un canvi radical en la classificació arxivística. Un dels més destacats és Greg Bak, molt crític amb la classificació funcional predominant en aquell moment, el qual proposava superar la regla d'un sol codi de classificació explotant al màxim la possibilitat de gestionar els documents en múltiples contextos mitjançant metadades (Bak 2012).

El propi Greg Bak plantejava una segona dimensió en la seva crítica. Des de fa temps, les classificacions unívokes s'estan abandonant (Bak 2012).¹⁰⁵ No només en els àmbits científics o professionals, sinó també en l'experiència d'usuari, on cada persona prefereix crear la seva classificació abans que seguir una classificació "oficial" preestablerta.

L'arxivística, malgrat l'evolució viscuda durant les darreres dècades, manté encara una visió de la classificació directament hereva dels autors de mitjans del segle XX (Bak 2012). La idea de Jenkinson que els arxius són una acumulació natural, espontània, d'evidències d'activitat segueix essent la línia directriu de les metodologies arxivístiques (Jenkinson 1922). Aquesta idea té un rerefons en la voluntat de situar l'arxivística com a ciència de ple dret, equiparable a les ciències naturals, a través de la identificació de lleis o regles absolutes.

En els arxius de dades l'accés i la descoberta de nova informació han passat a un primer pla, i s'han convertit en l'eix del tractament. Però no ho han fet a través dels tradicionals mètodes de classificació i descripció, sinó a través de tecnologies de *data mining* i processament de llenguatge natural amb *machine learning* i intel·ligència artificial (Mordell, 2022). Els instruments de descripció arxivístics clàssics, malgrat ser aplicables a les dades als efectes de governança, han deixat ja de ser (i potser definitivament) instruments de recuperació i accés a la informació.

Ja a finals dels 60 James Rhoads plantejava la previsible obsolescència dels instruments arxivístics tradicionals, en el que denominava "*cybernetic approach*", quan considerava que a través de les en aquell moment emergents bases de dades "*all the information currently known about the archives' records*" – including finding aids, lists, and catalogues – could be compiled to generate a special topic-based guide on the fly from user-provided search terms" (Rhoads, 1969). Ja es preconfigurava la possibilitat de superar el valor de les agrupacions documentals com a recurs de cerca a través de classificacions basades en el contingut de la informació descrita, en la línia de les actuals ontologies aplicables a documents d'arxiu. L'evolució de la tecnologia no ha deixat d'incrementar aquesta possibilitat (Yeo, 2012; Bell, 2020).

Segons Devon Mordell, "*For archives as data, access and discovery are ultimately emphasized. But not through the traditional methods of arrangement and description.*" (Mordell, 2022). Quan els arxius s'han "dadificat", es fa possible aplicar-hi tècniques de recuperació i anàlisi de la informació que fan gairebé innecessàries les classificacions basades en categories estanques. I l'ús de la intel·ligència artificial obre la porta a la generació automàtica de classificacions i al replantejament dels criteris més pertinents per classificar documents d'arxiu (Rolan et al., 2019).

¹⁰⁵ Podem citar com exemple la classificació biològica, en la qual s'estan deixant de classificar els organismes a partir de les seves característiques externes en favor d'una classificació a partir de la seva descendència evolutiva.

Un segon discurs crític amb les classificacions tradicionals s'ha obert també, tant en l'àmbit de l'arxivística com de la ciència de dades, arran de l'augment en la potència del tractament d'informació. L'automatització en l'assignació de categories o classes, tant si és basada en regles com en machine *learning*, i la possibilitat de desencadenar accions automàticament a partir d'aquestes classificacions, ha despertat el debat sobre els biaixos ideològics implícits en moltes classificacions (Bowker & Star, 1999; Boyd & Crawford, 2012). Terry Cook, quan apunta el concepte de reinvençió o construcció del passat a partir de la interpretació dels arxius, accepta implícitament que la subjectivitat és present en el tractament del material d'arxiu (Cook, 2013). Però aquest biaix és gairebé implícit en qualsevol estructura classificatòria, també existeix en les ontologies de la Web semàntica, i fins i tot en els modelats de les bases de dades (Yeo, 2012). Això no obstant, l'ús de la intel·ligència artificial, pel seu caràcter sovint no supervisat, amplifica notablement els riscos ètics dels biaixos en la classificació (Munn, 2022).

2.3.3.2 Crítica del criteri funcional en la classificació arxivística

Malgrat l'evolució en la tecnologia, i en els plantejaments sobre l'estructuració de la informació a totes les disciplines adjacents (Documentació, Ciència de dades, etc.), la pràctica arxivística segueix majoritàriament captiva d'un model de classificació funcional definit durant el darrer quart del segle passat (Bak, 2012). Però no només en l'aspecte pràctic: la teoria arxivística, incloent la que es fixa com a doctrina a través de les normes tècniques, segueix proposant com a criteri de classificació el criteri exclusivament funcional.

Una part de la crítica a aquest criteri no és substantiva, sinó que prové d'una aplicació negligent o poc curosa del criteri funcional en moltes organitzacions, i d'una manca de sistemàtica en l'elaboració dels quadres de classificació (Foscarini, 2012). Però també d'una manca de reflexió sobre la utilitat de la classificació arxivística (Franks, 2013; Mokhtar & Yusof, 2015). Les darreres aportacions metodològiques d'una certa rellevància son anteriors a 2010, i les crítiques teòriques publicades a partir d'aquesta data no s'han materialitzat posteriorment en propostes metodològiques alternatives. Així doncs, una de les aportacions de referència, el mètode australià DIRKS, data del 2001 (*DIRKS: A Strategic Approach to Managing Business Information*, 2001). Aquest mètode, amb una gran influència sobre el desenvolupament de la norma ISO 15489 (Healy, 2001; Macintosh & Real, 2007), i que fou la base del desenvolupament de la norma ISO 26122 per a l'anàlisi de funcions i activitats (*ISO/TR 26122:2008 Information and documentation — Work process analysis for records*, 2008), va quedar fora de context amb la segona versió de la norma ISO 15489, fortament criticada per la comunitat arxivística per la seva orientació exageradament tecnològica (Alberch, 2016; Findlay, 2018). D'altres propostes realment interessants, com la dels Arxius Nacionals del Regne Unit (Todd, 2003) o la proposta canadenca de l'estat l'Alberta (*Alberta's*

modified functional classification system model, 2007) tampoc han tingut després major impacte sobre la comunitat arxivística.

Alguns autors crítics amb la classificació funcional també argumenten que no es poden separar fàcilment les funcions dels òrgans que les desenvolupen, malgrat s'ha volgut considerar sovint la classificació funcional com oposada a la classificació orgànica (Foscarini, 2012).

Una de les oportunitats que s'ha plantejat de forma recurrent ha estat la possibilitat de fer coincidir la identificació de funcions i activitats amb el mapatge de processos (Alonso & Lloveras, 2007; Escobar, 2009; Sierra, 2012), aprofitant d'aquesta forma el *know-how* metodològic acumulat per la reenginyeria de processos.¹⁰⁶ Aquesta opció apareixia per primer cop en la norma ISO 26122 i en la seva doble proposta d'anàlisi funcional i anàlisi seqüencial, i fou recollida també en la distinció entre *Business Classification Scheme* i *Records Classification Scheme* que apuntava MoReq2010 (Vieira et al., 2012). Llevat d'algunes experiències puntuals (Serra & Casellas, 2007), aquesta no ha estat una línia determinant en els models arxivístics: els principals models no han incorporat el procés com l'entitat funcional principal, a partir de la qual s'identifiquen les entitats documentals. Com s'ha analitzat anteriorment, alguns models no esmenten el concepte de procés com a tal (e-EMGDE o RiC), mentre que d'altres l'assimilen a algun dels nivells de la tríada funció-activitat-transacció, preferentment al d'activitat (ISDF o NEDA).

Un exemple d'aquesta desconexió és el *Catàleg Corporatiu de Processos* de la Generalitat de Catalunya, l'eina que registra tots els processos i tràmits que realitzen els ens de la Generalitat.¹⁰⁷ Aquest sistema, gestionat actualment per la Direcció General d'Administració Digital, va incloure en la descripció de cada procés la correspondència amb l'entrada equivalent del *Quadre de Classificació Corporatiu* de la Subdirecció General d'Arxius de la Generalitat de Catalunya.¹⁰⁸ De forma coetània es va posar en marxa el desenvolupament de la nova versió del quadre de classificació corporatiu. I malgrat que en la guia metodològica del quadre s'indica que es va aplicar la norma ISO 26122 (*Guia de classificació de documents del sistema general de gestió de la documentació administrativa*, 2015), el resultat, publicat com annex a la *Instrucció 1/2015 de 26 de febrer de 2015, per la qual s'aprova l'instrument tècnic Quadre de classificació funcional del Protocol de Gestió de Documents Electrònics i Arxiu de la Generalitat de Catalunya i s'estableixen directrius per a la seva gestió i actualització*, i

¹⁰⁶ Vegeu introducció **GMAP** i **GDAD** (apartat 3.1 del capítol 3).

¹⁰⁷ Vegeu: <https://ctti.gencat.cat/ca/ctti/solucions-corporatives/gestio-de-recursos/gestio-informacio-i-documentacio/ccp/index.html>.

¹⁰⁸ Vegeu la presentació del projecte ja en producció a la II Jornada d'Innovació: <https://www.youtube.com/watch?v=1Ni-68ifkGs>.

actualment disponible en forma de dades obertes,¹⁰⁹ no manté un vincle conceptual determinant amb el *Catàleg Corporatiu de Processos*.¹¹⁰

Com ja hem comentat, un dels elements que ha propiciat la crítica al model funcional no ha estat només la seva manca d'evolució, sinó també la seva aplicació sovint mancada de sistemàtica (Mokhtar & Yusof, 2015). Moltes entitats, especialment del sector públic, segueixen mantenint quadres de classificació mono-jeràrquics i funcionals, però amb nombroses concessions orgàniques o a tipologies documentals, i sense incorporar una perspectiva orientada a processos.

Una part de la crítica a la classificació arxivística ve també dels seus usuaris (Ifould & Joseph, 2016; Chandler, 2022). La gestió documental va optar pel criteri funcional com una mesura de longevitat, a fi de fer més estable la classificació enfront dels canvis orgànics o de context (Sabourin, 2001). Però el criteri funcional, en tant que representació abstracta de l'activitat de l'organització, esdevé difícil d'entendre i aplicar per a molts usuaris, que es troben més còmodes amb classificacions menys precises però més intuïtives (Orr, 2005; Henttonen & Kettunen, 2011; Gunnlaugsdottir, 2012a, 2012b; Foscarini, 2012; Packalén, 2015; Bowker & Villamizar, 2017). Fins i tot els sistemes automatitzats d'auto-classificació de documents tenen dificultats quan han d'aplicar un criteri de classificació funcional (Payne, 2022).

Els usuaris finals es poden sentir més còmodes amb agrupacions que ajudin a la prioritització de les tasques, al desenvolupament dels projectes, o a la identificació ràpida per tipologies documentals (Yeo, 2012; Pan, 2017). De fet, alguns autors han proposat retornar a classificacions híbrides, matisant el caràcter funcional amb elements orgànics que facilitin l'aplicació als usuaris (Foscarini, 2012).

2.3.4 L'utilitarisme extrem i els intents de simplificació: els *Big Buckets*

La visió utilitarista que s'exposa en aquest apartat es pot considerar una conseqüència de les crítiques al model funcional comentades en l'apartat anterior. Visió a la qual considerem que la bibliografia acadèmica probablement no li ha prestat prou atenció, tenint en compte que planteja una alternativa radical a la classificació funcional, i especialment que desconstrueix totalment el concepte de sèrie documental emanada d'una activitat concreta.

¹⁰⁹ Vegeu: <https://analisi.transparenciacatalunya.cat/Sector-P-blic/Quadre-de-Classificaci-funcional-de-la-documentaci/5t23-dy8y>.

¹¹⁰ Més enllà d'una referència creuada en la implementació tecnològica del *Catàleg Corporatiu de Processos* (camp "Classificació").

El concepte utilitarista, ja descrit a l'apartat 2.3.1.2, conclou que la sèrie documental és una agrupació convencional orientada a poder aplicar un repertori de polítiques sobre els documents, que bàsicament en determinen la conservació, l'accés i l'ús. Per aquest motiu, sovint el nivell d'identificació de sèries documentals, més enllà del seu context natural de creació, s'ajusta a la forma concreta de les evidències que la componen, i en funció d'això als diferents tipus de regles de conservació, accés i ús que calgui aplicar. Aquest ús consuetudinari de la sèrie en la gestió documental ha provocat que la utilitat i els beneficis a nivell pràctic se sobreposin progressivament a la puresa del context de creació, i ha tingut un efecte pervers sobre la qualitat de les classificacions arxivístiques.¹¹¹ En molts casos, l'aplicació pràctica ha portat a reconèixer com a sèries conjunts d'evidències només pel seu vincle material (ubicades en un mateix lloc o en un mateix format), tecnològic (integrades dins un mateix sistema d'informació) o de facilitat d'ús (arxivables físicament en un mateix contenidor) (Foscarini, 2012; Mokhtar & Yusof, 2015; Ladley, 2016). Sovint s'ha acabat consolidant aquesta distorsió mitjançant quadres de classificació poc coherents, i fins i tot amb justificacions metodològiques (per exemple relatives a l'ús de sèries factícies o temàtiques en la gestió documental). El baix nivell de qualitat i coherència de les classificacions de molts sistemes de gestió documental actuals es deu a la prioritització d'aspectes pràctics o utilitaris per sobre del context original (Foscarini, 2012).

Cap a la dècada del 2010 la gestió documental de l'entorn anglosaxó, i especialment l'àmbit empresarial, va portar a l'extrem el concepte utilitarista de sèrie documental, arribant gairebé a obviar el component funcional i agrupant directament la documentació pels tipus de regles de conservació i accés, entenent aquests tipus bàsicament com la combinació de disparadors i terminis. Fou el que s'anomenà la teoria dels *Big Buckets* (Cisco, 2008a, 2008b; Galloway, 2008). La reminiscència funcional es va conservar només a alt nivell, al nivell del que s'anomenava *Big functions*.

L'arrel de la proposta la trobem en la dificultat abans esmentada de classificar adequadament i de forma sistemàtica tots els documents que es produeixen en un entorn digital per part del personal d'oficina. L'esforç per aplicar classificacions funcionals complexes i metadescriure manualment documents originava buits en la classificació que limitaven l'efecte de les polítiques de conservació i accés (Galloway, 2008). Davant aquesta evidència, Susan Cisco va proposar l'any 2007, inspirant-se en la solució de les *General Records Schedules* (GRS) aplicada per alguns arxius nacionals,¹¹² l'agrupació dels documents en grans categories, molt genèriques, vinculades a regles de conservació d'ampli abast (Cisco, 2008a, 2008b). Aquestes grans

¹¹¹ Alguns autors, com James Lappin, defineixen directament com a fracàs l'ús de classificacions arxivístiques sobre els documents electrònics, i especialment sobre el correu electrònic (Lappin, 2020b).

¹¹² Especialment pels Arxius Nacionals dels Estats Units (NARA), regularment actualitzades a <https://www.archives.gov/records-mgmt/grs>, i del Canadà (Suderman, 1997; Sabourin, 2001).

agrupacions no s'aplicarien als documents de conservació permanent, sinó només a aquelles sèries temporals (amb terminis de conservació inferiors als 20 anys), que es calculava que representaven el 90% del total de sèries de la majoria d'organitzacions (Galloway, 2008). El raonament era: a menys categories, menys probabilitat d'error per part del personal classificador. Ja en aquell moment alguns autors consideraven que la única solució viable era l'automatització, descartant la simplificació de les classificacions (Bailey, 2009). Malgrat nombroses limitacions metodològiques, el model dels *Big Buckets* ja plantejava la inclusió de qualsevol manifestació d'informació com a part dels continguts a recollir en els *buckets*, fossin documents o dades, amb l'objectiu de donar cobertura mitjançant polítiques a tota la informació generada per l'organització, i no només als documents formalitzats com a tals (Cisco, 2008b). La idea dels *big buckets*, malgrat no haver tingut un impacte destacat en el món arxivístic, continua vigent en noves formulacions realitzades a redós del l'emergència dels entorns de treball col·laboratiu, on les decisions de l'usuari final són determinants per a la gestió dels documents (Lappin, 2020b, 2021; Skjekkeland, 2021).

El model *Capstone* és un altre exemple d'aquesta voluntat simplificadora, en aquest cas aplicada al correu electrònic. L'origen es troba en un memoràndum del President Barack Obama del 28 de novembre de 2011 sobre la gestió dels documents del Govern,¹¹³ i en la directiva emesa posteriorment, el 24 d'agost de 2012, per l'OMB i el NARA,¹¹⁴ on s'especificava “*By December 31, 2016, federal agencies must manage all email records in an electronic format*”, i més endavant “*By December 31, 2017, [...] will make substantive changes to the General Records Schedules (GRS). These significant changes will include, combining the records series into more appropriate aggregations for easier disposition action by agencies, and expanding the number of permanent records series in the GRS to reduce the scheduling and appraisal burden on agencies.*”. L'any 2013 el NARA va publicar l'anomenat butlletí *Capstone*, titulat *Guidance on a New Approach to Managing Email Records*.¹¹⁵ El canvi que es proposava anava més enllà del fet de deixar d'imprimir correus electrònics, i afectava directament la classificació arxivística. La instrucció era deixar de classificar el correu electrònic en funció de l'activitat a la qual feia referència cada missatge (deixar de classificar per sèries, o en tot cas classificar per agrupacions de sèries), i passar a gestionar la bústia de correu com un bloc, aplicant polítiques arxivístiques a la bústia en funció del càrrec o rol del

¹¹³ Obama, B. (2011). *Presidential Memorandum: Managing Government Records*. The White House: Office of the Press Secretary. <https://obamawhitehouse.archives.gov/the-press-office/2011/11/28/presidential-memorandum-managing-government-records>.

¹¹⁴ *Managing Government Records Directive (M-12-18)*. (2012). Office of Management and Budget (OMB). National Archives and Records Administration (NARA).

¹¹⁵ *Guidance on a New Approach to Managing Email Records (Bulletin 2013-02)*. (2013). National Archives and Records Administration (NARA). Federal Records Management.

seu propietari.¹¹⁶ L'any 2015 el NARA va publicar la regla de conservació *GRS 6.1: Email managed under a capstone approach*,¹¹⁷ determinant els terminis de conservació del correu electrònic per a les agències federals que apliquen el mètode *Capstone*.¹¹⁸

L'esperit del mètode *Capstone* també era present en el món privat, on l'arxivat en sèries documentals dels correus electrònics als efectes de compliance es considerava també una tasca cada cop més difícil de portar al dia, atès el nombre creixent de missatgeria i la sobrecàrrega de treball dels gestors (Saffady, 2014). William Saffady proposava l'ús de *Uniform retention rules*, és a dir terminis fixos per a la conservació del correu, que oscil·laven entre un mínim de 3 anys i un màxim de 10, i resolució d'eliminació total del correu, llevat de les excepcions que fossin indispensables per motius legals. Com més llarg el termini, menys excepcions a gestionar, i per tant menor risc per l'aplicació d'una regla uniforme (Saffady, 2014). Aquesta perspectiva és considerada tanmateix excessivament simplificadora, atès que els missatges que s'han de conservar com excepcions, aproximadament un 10% segons James Lappin, són habitualment els que concentren el major valor evidencial de tot compte de correu (Lappin, 2020b).

El mètode *Capstone* s'ha seguit aplicant de forma regular, i recentment ha ampliat el seu abast. En primer lloc amb la promulgació de la *Electronic Message Preservation Act* (ECPA).¹¹⁹ I en segon lloc per la publicació per part del NARA d'una ampliació de l'abast del mètode *Capstone*, que inclou la gestió i preservació de la missatgeria electrònica seguint el mateix patró que el correu electrònic.¹²⁰ L'impacte d'aquesta directriu pot tenir una repercussió molt important en la gestió documental electrònica, especialment amb la transformació que s'està operant amb el model de treball amb els

¹¹⁶ El butlletí *Capstone* diu literalment: “*Capstone offers agencies the option of using a more simplified and automated approach to managing email, as opposed to using either print and file systems or records management applications that require staff to file email records individually. Using this approach, an agency can categorize and schedule email based on the work and/or position of the email account owner.*”

¹¹⁷ *White Paper on The Capstone Approach and Capstone GRS*. (2015). National Archives and Records Administration (NARA). <https://www.archives.gov/files/records-mgmt/email-management/final-capstone-white-paper.pdf>.

¹¹⁸ L'Administració Federal dels Estats Units aplica un mètode d'avaluació i tria descentralitzat. Els Arxius Nacionals (NARA) publiquen regles que defineixen el marc (*General Records Schedules*), però que no baixen al nivell de sèrie documental. Aquesta vinculació del marc general amb la sèrie documental té lloc a cada organització, que elabora les *Specific Records Schedules*. Aquesta estructura concedeix a cada ens un cert marge d'autonomia per aplicar d'una forma més o menys estricta les GRS. Vegeu: <https://www.archives.gov/records-mgmt/grs>.

¹¹⁹ *H.R.1582 - 116th Congress (2019-2020): Electronic Message Preservation Act*. (2019, març 13). <https://www.congress.gov/bill/116th-congress/house-bill/1582/text>.

¹²⁰ *Expanding the Use of a Role-Based Approach (Capstone) for Electronic Messages (Bulletin 2023-02)*. (2023). National Archives and Records Administration (NARA). Federal Records Management. <https://www.archives.gov/records-mgmt/bulletins/2023/2023-02>.

documents i dades en els entorns de col·laboració integrats, i concretament amb Microsoft TEAMS (Lappin, 2020b, 2021).

2.3.5 La introducció dels models ontològics en la classificació i la descripció arxivístiques

En paral·lel als plantejaments crítics amb les classificacions jeràrquiques i el criteri funcional, aparegué fa una dècada un corrent que propugnava l'ús de plantejaments ontològics per a la classificació i descripció de recursos d'informació en l'àmbit dels arxius. Un dels primers autors a desenvolupar aquesta possibilitat fou Ralf Klischewski. Partint de la constatació de l'ús de tesaurus per a la classificació de documentació administrativa en diferents contextos nacionals (Austràlia, Regne Unit, Canada o Suïssa), Klischewski va aplicar un plantejament ontològic per a la millora del sistema de gestió documental de l'Administració regional del *land* de Schleswig-Holstein, en un projecte anomenat *ParoPlus* (Klischewski, 2006). Denise Bedford va apuntar la mateixa idea en una conferència titulada "*Records Management Ontologies*", presentada al *Collaborative Expedition Workshop* l'any 2007.¹²¹ Jaffar Alalwan i Manof Thomas elaboraren una proposta de base ontològica per a l'avaluació de sistemes de gestió documental (Alalwan & Thomas, 2011, 2012). En aquest moment les biblioteques ja treballaven des de feina temps en la incorporació d'ontologies a la descripció de recursos, amb projectes com Europeana,¹²² i els arxius s'hi sumaven amb projectes com Archives Hub¹²³ o una primera ontologia basada en la norma *Encoded Archival Context* (EAC) de la *Library of Congress* (Mazzini & Ricci, 2011). Una iniciativa pionera a l'estat espanyol fou el model multidimensional de classificació que Alejandro Delgado va aplicar a l'Arxiu Municipal de Cartagena (Delgado, 2010). També fou pioner, en l'àmbit normatiu, el projecte que va impulsar a Espanya el *Ministerio de Cultura* amb la creació de la *Comisión de las Normas Españolas de Descripción Archivística* (CNEDA)¹²⁴ (Moyano, 2013), la qual esdevingué un precedent directe per a l'elaboració de la norma de descripció arxivística *Records in Context* del Consell Internacional d'Arxius (Gueguen et al., 2013; Llanes-Padrón & Pastor-Sánchez, 2017). I en l'àmbit català, els treballs de Julio Quílez han encetat també una línia de treball en ontologies arxivístiques (Quílez, 2011, 2017). En l'àmbit de la preservació digital, el projecte PERICLES, desenvolupat

¹²¹ Presentació disponible a <https://slideplayer.com/slide/697721/>.

¹²² Vegeu <https://www.europeana.eu/es>.

¹²³ Vegeu <https://archiveshub.jisc.ac.uk/>.

¹²⁴ Vegeu referències al capítol 4 d'aquesta tesi.

dins el 7è programa marc de la Unió Europea, va plantejar també una ontologia per a la preservació digital.¹²⁵

Un dels escenaris on les ontologies tenen un potencial major d'implantació és en la governança de les dades, dins l'àmbit del ja esmentat *data discovery* (Wang, 2014; Rajmohan et al., 2023). Les ontologies se situen en la capa d'exploració d'un *data lake*, i permeten gestionar la provenença de les dades i estructurar els catàlegs de dades, i especialment els *business glossaries* associats a aquests catàlegs (Kononow, 2019; Dibowski et al., 2020). L'ús d'ontologies implica sovint una graficació de les relacions, en forma de *knowledge graphs* (Schmid et al., 2019; Smajevic et al., 2021; Smajevic & Bork, 2021; Chessa et al., 2022; Li et al., 2022), combinació que esdevé un element cabdal per a la governança de les dades en els sistemes distribuïts (Li et al., 2022).

El gran pas en la introducció dels plantejaments ontològics en els arxius ha estat sens dubte la publicació de la norma de descripció arxivística "*Records in Context*", amb la qual el Consell Internacional d'Arxius té previst substituir la família de normes ISAD(G) (Llanes-Padrón & Pastor-Sánchez, 2017; Clavaud & Wildi, 2021). La norma RiC encara està en un estadi incipient, i probablement la seva aplicació permetrà millorar algunes febleses, com el fet de no tenir en compte el mapatge amb altres ontologies amb les quals guarda relacions evidents, com l'ontologia PROV-O¹²⁶ (Hartig & Zhao, 2010; Duranti, 2016). En tot cas, RiC marca la tendència de futur de la descripció arxivística: nombrosos projectes s'adrecen a garantir la interoperabilitat de la resta de normes amb RiC (Francart et al., 2021; Garmendia & Retter, 2021), i alguns ja plantegen basar sistemes complets de gestió documental i arxiu en plantejaments ontològics (Aguayo, 2021; Costa & Ramalho, 2021; Hawkins, 2021). La mancança que apuntava Fiorella Foscarini fa una dècada dona indicis doncs d'estar-se revertint.¹²⁷

2.3.6 La dadificació de la sèrie documental

El concepte de conjunt de dades estructurades, assimilat a base de dades, apareix en l'arxivística com una nova forma simbòlica (Yeo, 2012), que planteja una narrativa diferent dels fets documentats (Bowker & Star, 1999; Yeo, 2012). Així ho van tractar els primers projectes d'arxiu de bases de dades, el pioner dels quals fou el *National Digital Archive of Datasets* (NDAD),¹²⁸ dels Arxius Nacionals del Regne Unit. I així ho han tractat

¹²⁵ Vegeu <https://cordis.europa.eu/project/id/601138>. El projecte, actualment tancat, no permet l'accés a alguns dels continguts publicats en el seu moment.

¹²⁶ Vegeu <https://www.w3.org/TR/prov-o/>.

¹²⁷ Foscarini afirmava que la visió relacional de les funcions i activitats, i sobretot l'anàlisi de les relacions entre processos, era encara lluny de les pràctiques i els coneixements dels arxiviers (Foscarini, 2012).

¹²⁸ Vegeu <https://discovery.nationalarchives.gov.uk/details/r/C18822>.

els posteriors projectes de preservació de bases de dades, com SIARD¹²⁹ als Arxius Nacionals de Suïssa (Heuscher et al., 2004), o RODA als Arxius Nacionals de Portugal (Ramalho, José Carlos et al., 2007). Els conjunts de dades es consideraven una nova forma documental, que es podia descriure, igual que qualsevol altra manifestació documental, amb la norma ISAD(G) (Shepherd & Smith, 2000).

En la literatura arxivística no abunden els intents de dotar de significació arxivística els conjunts de dades, i s'han acceptat acríticament com una nova font documental a preservar (Niu, 2016). En algun cas s'ha fet intents de discriminar els conjunts de dades en funció del seu origen, però no s'ha arribat a generar doctrina. Campos fa una aportació interessant en aquest sentit, distingint entre els conjunts de dades de l'organització, aquells que poden ser susceptibles d'equiparació a agrupacions documentals, les dades-producte, és a dir les creades a partir de dades ja existents en els arxius, i les dades-donació, provinents majoritàriament de fons privats o de recerca (Campos, 2023).

Però els sistemes d'informació no processen objectes, processen dades. En conseqüència, per a poder processar la informació d'un arxiu, i especialment si es pretén fer-ne analítica, cal que estigui en forma de dades, i les dades en sí seran l'objecte de l'atenció arxivística (Padilla, 2017).¹³⁰ A aquest fet cal afegir l'extensió del concepte d'evidència a qualsevol contingut d'un sistema d'informació, el que anomenem "recordness" d'un sistema, un concepte proposat per Geoffrey Yeo (Yeo, 2007) i present a l'estudi de Borglund i Engvall (Borglund & Engvall, 2014) per definir un sistema que conté dades autèntiques que mai no han existit en forma de document.

El concepte de dadificació no és nou en l'àmbit de la governança de les dades i el *Big data*, i s'ha definit com "*the transformation of social action into online quantified data, thus allowing for real-time tracking and predictive analysis*" (Mayer-Schönberger & Cukier, 2014). Aquest concepte de dadificació, aplicat als documents d'arxiu, malgrat no estar consolidat doctrinalment, apareix també cada cop amb més freqüència en la bibliografia arxivística, on es comença a utilitzar el concepte "*archives as data*" (Mordell, 2022).

El concepte de dadificació, i el d'expedient basat en dades, també era present des dels inicis en els sistemes de *Workflow* i BPM (Van Der Aalst et al., 2005). Aalst, Weske i Grünbauer consideraven que el *case* (és a dir l'expedient) pot ser considerat un

¹²⁹ SIARD segueix el model OAIIS de preservació digital (vegeu <https://public.ccsds.org/pubs/650x0m2.pdf>), i permet convertir bases de dades relacionals d'ús habitual (Microsoft Access, DB/2, MySQL, Oracle, PostgreSQL o SQL Server) a un format obert que en garanteix la preservació, el format SIARD. Per a més informació vegeu <https://www.bar.admin.ch/bar/en/home/archiving/tools/siard-suite.html>.

¹³⁰ Thomas Padilla afirma "*to see collections as data begins with reframing all digital objects as data*".

“producte”, amb una estructura i un estat, i que aquesta estructura està basada en el que anomenaven “*data objects*”, els quals definien com:

“[...] a piece of information which is present or not present and when it is present is has a value.”

També destacaven que, en tot moment, el responsable de l'expedient ha de tenir accés a totes les dades de l'expedient, i que els formularis eren formes de presentació agregada i temporal de les dades de l'expedient (Van Der Aalst et al., 2005). I en la mateixa línia, Mutschler, Weber i Reichert parlaven de *data-driven case handling* (Mutschler et al., 2008).

Però amb independència de l'aparició del terme, el concepte de dadificació, entès com la conversió a dades d'un conjunt de documents, ja el trobem a principis dels 2000 en la metodologia APS (Puig-Pey & Guiu, 2002, 2008). El concepte de dadificació APS és tanmateix diferent del que apunten autors com Devon Mordell, on la dadificació s'ubica en l'àmbit dels arxius històrics i en el tractament del text de documents originalment en paper. L'anàlisi numèric de documents digitalitzats via OCR no és nou, com tampoc ho és el seu tractament i explotació amb eines de *machine learning*. Però cal diferenciar entre la digitalització de documents analògics, que produeix documents llegibles per màquina, i el concepte de dadificació, que implica que les dades que s'obtenen dels documents poden ser tabulades i analitzades (Mayer-Schönberger & Cukier, 2014; Mordell, 2019, 2022). Per aquest motiu, Devon Mordell afirma que “*the datafication of archives is not just a technical process but also a technocultural one*”, i proposa que l'objectiu actual de l'arxivística ha de ser “*renders the archive amenable to computational analysis*” (Mordell, 2022). Per aquesta via l'arxivística es connecta sovint a l'activisme de dades (Currie & Paris, 2018) i als corrents que promouen, especialment en el sector públic, la publicació de dades obertes (Chorley, 2017; Shepherd et al., 2019).

Segons Yeo, l'agrupació documental¹³¹ és lluny d'haver quedat obsoleta, però la seva forma de manifestar-se ha canviat radicalment (Yeo, 2012). Geoffrey Yeo argumenta que l'assimilació de document, sèrie i expedient com una jerarquia única, i el fet d'assimilar el concepte de “*record*” amb el d'unitat documental simple (“*item*”) són herències del món pre-digital (Yeo, 2012). Una visió més autònoma i flexible del concepte de “*record*”, com la que proposem en aquesta tesi amb la matriu de variables, ja l'apuntaven Sheperd i Yeo en el seu manual (Shepherd & Yeo, 2003).¹³² I compartien amb d'altres teòrics la visió que la sèrie documental existeix sense la necessitat d'una materialització física, només pel fet que es generen evidències d'un procés (Duranti et al., 2002; Yeo, 2012).

¹³¹ Yeo l'esmenta com a “col·lecció” (Yeo, 2012).

¹³² “*The records of the various steps can be aggregated to form a record representing the activity in total. [...] records exist at two levels: the elementary level of the individual steps and the higher level of the activity.*” (Yeo, 2012, p. 51)

Un element especialment rellevant a destacar és que la dadificació és la base de l'evolució actual en la forma de comunicar la informació, que és cada cop més visual (Pérez-Montoro, 2020). En una taula rodona virtual, *Recordkeeping theory, models & strategies and today's workplace*, organitzada per la *Recordkeeping Roundtable* el 18 de març de 2021, Chris Hurley afirmava que ja no és possible treballar amb estructures fixes, i que les estructures de consum de la informació han de ser dinàmiques i adaptades al moment i les necessitats de cada usuari, tot i que destacava que encara no hi ha teoria arxivística en aquest sentit (Findlay & Cumming, 2021).

El model de visualització interactiva actual té els seus orígens en el concepte de quadre de comandament integral (Kaplan & Norton, 1996), que els seus autors ja van definir com un "sistema de gestió". La seva progressió ha estat lenta, a causa de la complexitat per a construir els quadres, i especialment per la necessitat de disposar d'un elevat nombre de fonts d'informació estructurades (dadificades) i integrades (Kandel et al., 2011). Però l'evolució de la tecnologia ha permès que les visualitzacions estiguin actualment a l'abast de l'usuari final a un cost molt raonable, fent de la visualització de dades un element cada cop més important a les organitzacions (Aldama et al., 2022).¹³³ Aquest accés transversal a les dades entronca directament amb el corrent de democratització de l'accés intern a les dades que inspira el model del *data mesh* (Dehghani, 2019; Delgado, 2022).

Alguns autors consideren tanmateix que usuaris sense coneixements d'analítica de dades no podran extreure tot el valor de les visualitzacions de dades (Lemieux et al., 2014; Lemieux, 2015), i que la visualització com a producte ha de seguir evolucionant amb serveis d'informació personalitzats i en temps real que utilitzin intel·ligència artificial (Marr, 2022).

Per tots aquests motius, en les conclusions de l'objectiu 6 apuntem la idea que la visualització de dades es pot arribar a considerar una nova forma documental, probablement amb importants repercussions en la teoria arxivística encara per analitzar.

¹³³ A dia d'avui l'usuari final disposa d'entorns que permeten generar visualitzacions avançades amb poc esforç i de forma amigable, com *Tableau* (<https://www.tableau.com/es-es>), *Google Looker Studio* (<https://lookerstudio.google.com/navigation/reporting>) o *Microsoft Power BI* (<https://powerbi.microsoft.com/es-es/>).

Capítol 3. Metodologia per a la construcció del sistema de governança

3.1 Introducció

El món de la consultoria en organització i qualitat ha identificat les diferents entitats que intervenen en la gestió de la informació d'una organització amb un marc conceptual i metodològic propi.¹³⁴ Així ho ha fet també la disciplina informàtica en l'anàlisi per al desenvolupament d'aplicacions, sovint lligat a lectures focals (Rittgen, 2006). I fins i tot la gestió documental ha prioritzat en molts casos les visions particulars de cada organització¹³⁵. Això ha provocat que el punt feble de la major part de metodologies d'anàlisi de processos, sèries documentals i aplicacions informàtiques sigui la identificació dels objectes des d'un punt de vista global, relacional, i no pas focal o sectorial, i en segon lloc la sistemàtica en la seva transformació cap a models integrats. Aquestes són les dues aportacions principals que es volen fer amb aquesta metodologia i les dues guies que la desenvolupen.

3.1.1 Objectius

L'objectiu d'aquesta metodologia és construir una base de coneixement sobre l'organització que pugui servir per a portar a terme una governança integrada de la informació.¹³⁶

La governança de la informació té una dimensió passiva, analítica, i una dimensió activa, transformadora. La dimensió analítica es basa en conèixer amb detall l'organització, i mantenir aquest coneixement al llarg del temps (el que s'anomena monitoritzar). La

¹³⁴ Trobem exemples en models metodològics com el model *Lean Six Sigma*, o model EFQM, consultable a <https://efqm.org/>. Vegeu per aquests dos models Pyzdek & Keller, 2018, i també Fonseca et al., 2021.

¹³⁵ Com demostra l'estesa afirmació que és impossible fer un quadre de classificació que serveixi per a més d'una organització, malgrat la pràctica ha demostrat el contrari. A títol d'exemple vegeu el *Protocol de suport als arxius d'entitats i associacions culturals*, on es planteja un quadre de classificació aplicable de forma directa a tot tipus d'entitats (https://xac.gencat.cat/web/.content/xac/00_home_xac/noticies/2020_suport_associacions/Protocol-de-suport-als-arxius-d'entitats-i-associacions-culturals.pdf).

¹³⁶ Malgrat l'aplicació de la metodologia en el marc d'aquesta tesi ha estat orientada principalment a la recollida, estructuració i anàlisi del cas d'aplicació, la metodologia s'ha desenvolupat amb una vocació global o orientada a facilitar una eina que pugui ser d'aplicació en altres contextos.

dimensió transformadora es basa en, partint del coneixement disponible, aplicar canvis coordinats en diferents elements del conjunt per assolir canvis en positiu en el conjunt de l'activitat de l'organització.

Amb aquesta finalitat s'ha dissenyat una metodologia que es condensa en dos instruments: la **Guia d'inventari i mapatge**, que identificarem amb les inicials **GMAP**, i la **Guia de dadificació de processos**, que identificarem amb les inicials **GDAD**.

L'objectiu de la **GMAP** és construir un mapa de les entitats existents en una organització que siguin determinants per a la gestió de la seva informació. S'utilitza el terme mapa, i no el de catàleg o inventari, perquè el que es pretén és:

- a) Conèixer l'existència de cada entitat, i identificar-la de forma objectiva i unívoca.
- b) Caracteritzar cada entitat (descriptiva i categòricament).
- c) Ubicar-la dins el conjunt de l'organització a partir de les seves relacions amb la resta d'entitats.

El plantejament que es proposa parteix d'una base ontològica. Tot i així, la voluntat d'aquesta metodologia no és la construcció d'una ontologia concreta, sinó posar les bases per a permetre la construcció d'ontologies a nivell d'una organització, el que algun autor anomena ontologia administrativa (Aguayo, 2021), i oferir una eina metodològica per identificar, caracteritzar i relacionar els elements que n'han de formar part.¹³⁷

L'estructura dels mapes està pensada per consolidar el coneixement obtingut, fent fàcil el seu manteniment i actualització, juntament amb l'anàlisi del seu contingut per a la identificació d'oportunitats de transformació i millora. Amb aquesta finalitat s'ha desenvolupat un conjunt d'indicadors d'anàlisi i seguiment que es descriuen al capítol 5 d'aquesta tesi.

Atès que una part important de les millores que s'identifiquen tenen relació amb la transformació digital dels processos, i que els guanys més evidents de la transformació s'obtenen quan el resultat són dades explotables, s'ha inclòs una segona guia, la **GDAD**. Aquesta guia proposa una sistemàtica per a, des de la perspectiva d'una correcta gestió documental, transformar tots els documents o evidències d'un procés en dades estructurades, i en conseqüència explotables.

Aquesta segona part de la metodologia té doncs com a objectiu la millora del substrat informacional dels processos, especialment del circuit pel qual es genera i consolida aquest substrat, i la seva posterior dadificació, per tal que totes les evidències generades estiguin basades en dades estructurades. Es contempla el procés complet, des de l'optimització d'un circuit documental amb mancances o disfuncions, fins a la transformació completa del circuit en un model basat en dades, i totalment integrat en el mapa de dades de l'organització. Tot allò que fa referència a l'optimització pura dels

¹³⁷ La codificació del model de dades en OWL, malgrat s'ha considerat fora de l'abast d'aquesta tesi, s'ha plantejat com una línia de treball a desenvolupar en el futur.

processos (temps d'execució, camí crític, colls d'ampolla, etc.) com a les eines que hi han de donar suport (desenvolupament d'aplicacions o solucions informàtiques) queda fora de l'abast d'aquesta metodologia.

Aquesta metodologia és tecnològicament neutra, i pot ser implementada en diferents tipus d'eines. En aquest projecte s'ha implementat en una eina desenvolupada específicament per al projecte, l'eina **CATPROC**, que es descriu en el capítol 4.

3.1.2 Bases metodològiques i models de referència

El marc conceptual en el qual s'incardina la metodologia que aquí es proposa és el de la governança de la informació a les organitzacions. Aquest és un concepte tremendament ampli i que s'ha abordat des de múltiples visions i disciplines. Per elaborar aquesta metodologia ens hem centrat preferentment en l'abordatge que s'ha fet des de l'arxivística i la gestió documental, i hem ampliat puntualment la visió cap a altres disciplines relacionades amb la gestió de la informació, en aquells aspectes que la gestió documental no havia plantejat. Aquesta aproximació s'ha materialitzat en la **GMAP**, amb la voluntat d'operativitzar aquesta governança de la informació en una organització concreta, i de forma centrada en les operacions de descripció, mapatge i monitorització dels canvis.

Per facilitar la seva lectura, no s'ha inclòs al redactat de les guies referències a bibliografia ni altres recursos d'informació. Totes les referències als models i estàndards utilitzats se citen en aquest apartat 3.1.2.

3.1.2.1 Models per a la construcció de sistemes de governança de la informació

El concepte de governança integral de la informació d'una organització, en tots els seus modelats i manifestacions, ha estat des de finals dels 90 una aspiració de la gestió documental (Upward, 2019; Soler Jiménez, 2020). Però els impulsos més rellevants en aquest camí no han vingut de l'arxivística acadèmica sinó del sector privat.

L'any 2014, EDRM, una comunitat de juristes i experts en *e-Discovery*,¹³⁸ va substituir en el seu model *E-Discovery Reference Model* (EDRM) el terme "*Records management*" pel de "*Information governance*", i poc després va publicar l'*Information Governance*

¹³⁸ Vegeu <https://edrm.net/>.

Reference Model (IGRM),¹³⁹ un model de governança integrat que incorpora la gestió documental com una de les grans àrees de la governança de la informació. És un plantejament que coincideix en gran mesura amb el model que va publicar l'any 2017 *DAMA International*¹⁴⁰ per a la governança integrada de la dada a les organitzacions, actualment desenvolupat en la segona edició del DMBOK (Data Management Association et al., 2017), i que també inclou la gestió documental com un element essencial del conjunt. I més recentment, *ARMA International*¹⁴¹ ha publicat una monografia dedicada exclusivament a la governança de la informació, el *Information Governance Body of Knowledge* (IGBOK),¹⁴² que s'afegeix al model de maduresa per a la governança de la informació basat en els coneguts *Generally Accepted Recordkeeping Principles* (GARP),¹⁴³ publicats per *ARMA International* l'any 2009 i actualitzats l'any 2017.

De fet, el terme gestió documental (“*document management*” o “*records management*”) i els termes que defineixen tecnologies derivades (“*Enterprise Content Management* (ECM)” o “*Content Services Platforms* (CSP)”) han estat progressivament substituïts en l'àmbit privat i tecnològic per termes que inclouen el concepte informació, com “*Records and Information Management* (RIM)”, “*Intelligent Information Management* (IIM)” o “*Information Lifecycle Management* (ILM)” (Kampffmeyer, 2023).¹⁴⁴

Aquesta tendència es va voler consolidar amb la publicació de la família de normes ISO 30300, establint la gestió documental com una funció d'alt nivell integrada amb la resta de funcions de l'organització (legal, tecnològica, etc.), clarament en la línia del model IGRM (Bustelo, 2011). Però no ha estat fins l'aparició l'any 2022 de la norma *ISO 24143:2022 Information and documentation — Information Governance — Concept and principles* que la governança integrada de la informació ha adquirit caràcter formal, obrint-se la porta al desenvolupament tant de normes més operatives com de metodologies i eines per donar-hi suport (Kampffmeyer, 2023). Malgrat la integració entre estàndards segueix essent una assignatura pendent, com evidencia la confirmació l'any 2022 de la norma *ISO/IEC 38505-1:2017 Information technology — Governance of IT — Governance of data — Part 1: Application of ISO/IEC 38500 to the governance of data* de forma independent de la ISO 24143, cada cop més es fa present en la disciplina

¹³⁹ Disponible a <https://edrm.net/resources/frameworks-and-standards/information-governance-reference-model/>.

¹⁴⁰ Vegeu <https://www.dama.org/cpages/home>.

¹⁴¹ L'associació internacional de professionals de la gestió documental, vegeu lloc Web a <https://www.arma.org/>.

¹⁴² Disponible a https://www.arma.org/page/Information_Governance.

¹⁴³ Disponibles a <https://www.arma.org/page/principles>.

¹⁴⁴ És il·lustratiu d'aquesta evolució observar els canvis produïts en la terminologia utilitzada per la *Association for Intelligent Information Management* (AIIM), disponible a <https://www.aiim.org/>.

la voluntat de conciliar els diferents àmbits de regulació tècnica (ISO/TC46/SC11-Archives/records management, 2023). La nostra aportació amb aquestes dues guies va justament en aquest sentit.

3.1.2.2 Models per al mapatge de processos i de sèries documentals

Nombrosos autors han destacat la dificultat de sistematitzar l'anàlisi de funcions i, per extensió, de processos (Sabourin, 2001; Foscarini, 2012; Casellas, 2019). Malgrat incidir reiteradament els autors en la seva importància (Harries, 2012), no abunden les metodologies que detallin, de forma operativa, com identificar i mapar els processos, i menys el seu substrat informacional, llevat d'alguns intents recents (Bell, 2020). Però hem trobat en les metodologies arxivístiques alguns referents útils.

DIRKS (*DIRKS: A Strategic Approach to Managing Business Information*, 2001) fou una de les primeres metodologies integrals per al disseny i implantació d'un sistema de gestió documental. DIRKS fou desenvolupat pels Arxius Nacionals d'Austràlia, juntament amb la *State Records Authority of New South Wales*. Contemporani de l'aparició i popularització en l'àmbit arxivístic de la norma ISO 15489 de gestió documental, també d'origen australià, DIRKS va esdevenir un referent per a la sistematització del disseny d'un sistema de gestió documental (Macintosh & Real, 2007). L'any 2007 DIRKS fou substituït per altres guies produïdes pels Arxius Nacionals d'Austràlia.

DIRKS plantejava la creació dels tres elements bàsics que la norma ISO 15489 estableix com obligatoris en tot sistema de gestió documental: un quadre de classificació de funcions, activitats i transaccions, un conjunt de regles de conservació, i un conjunt de regles de seguretat i accés, juntament amb l'esquema de metadades que hi doni suport. Per a la construcció d'aquests tres elements i la seva automatització, implantació i avaluació, DIRKS proposava una seqüència d'accions en vuit passos (del pas A al pas H). Bona part de les aportacions de DIRKS han quedat absorbides per normes posteriors, i tota la definició de requeriments, automatització i implantació ha quedat coberta per la família de normes ISO 16175 (ISO 16175-1:2010 – Part 1, ISO 16175-2:2011 – Part 2 i ISO 16175-3:2010 – Part 3). Però els passos A i B encara són metodològicament vigents en els sistemes de classificació arxivístics actuals. En la nostra metodologia hem tingut en compte el pas *A: Preliminary investigation* per sistematitzar el treball de camp i la recollida d'informació de diferents fonts, i especialment el pas *B: Analysis of business activity* per a l'anàlisi de les funcions i activitats i la seva transformació en una estructura de classificació funcional.

L'enfocament a processos del mapatge de funcions, activitats i sèries documentals, i l'anàlisi a dos nivells, el funcional i el seqüencial, es descriuen detalladament a la norma

ISO/TR 26122:2008 Information and documentation — Work process analysis for records. Aquesta norma desenvolupa el requeriment de la primera versió de la ISO 15489, que fixa com a primer pas de la norma la construcció d'un quadre de classificació de funcions i activitats, i té molts punts de contacte amb DIRKS. Malauradament ha estat una norma infrautilitzada en l'àmbit arxivístic i de la gestió documental, i l'omissió que feu la darrera versió de la ISO 15489 de l'ús de metodologies sistemàtiques encara la deixa més fora de context (Convery, 2016).¹⁴⁵

L'enfocament a processos de la classificació arxivística ha estat repetidament incorporat a la gestió documental, en alguns casos vinculat directament a models propis de la gestió de la qualitat, com EFQM (Alonso & Lloveras, 2007), i en d'altres a models més vinculats a la gestió documental i l'arxivística (Escobar, 2009; Sierra, 2012; Casellas, 2014; Mokhtar et al., 2016; Bell, 2020), però sense arribar a quallar en una formulació sistemàtica, més enllà dels models anteriorment citats. En aquest cas, les experiències prèvies de l'autor també s'han incorporat a l'enfocament d'aquestes guies (Serra & Casellas, 2007), juntament amb l'experiència de l'Administració de Portugal, especialment avançada en el concepte de mapa de sèries (Lourenço et al., 2020; Penteado, 2015, 2018).

3.1.2.3 Models per al mapatge de sistemes d'informació i aplicacions

Considerar els sistemes d'informació (en sentit ampli) i les aplicacions informàtiques (en sentit estricte) com un actiu d'informació, i per tant com a part d'un sistema integral de governança de la informació, ha estat tradicionalment un dels punts febles dels professionals de la gestió documental. Sigui per manca de formació tecnològica, sigui per manca de competència al nivell orgànic, el fet és que els sistemes d'informació s'han considerat més una eina o agent que no pas un actiu d'informació, com evidencia l'anàlisi realitzada en els punts 4.5, 4.6.1.1.1 i 4.6.1.11.1 del model de dades (capítol 4).

Pocs són doncs els exemples de pautes arxivístiques per a l'inventari i gestió dels sistemes d'informació. Novament el referent en aquest sentit és l'arxivística australiana. La idea de considerar els sistemes d'informació un element més del sistema de gestió documental la trobem ja en la norma AS 4390, precedent de la ISO 15489, i en els passos C a H del mètode DIRKS. Ho trobem novament en la família de normes ISO 16175. Però l'aportació més rellevant la trobem en el recent marc per a la governança

¹⁴⁵ Probablement la qualitat i coherència de molts quadres de classificació funcionals seria més alta si aquesta norma hagués estat més popular (Foscarini, 2009; Bustelo, 2012, 2017; Guardado, 2013; Packalén, 2015).

de la informació dels Arxius Nacionals d' Austràlia,¹⁴⁶ vinculats al qual trobem l'*Information Management Standard for Australian Government*, un estàndard basat en vuit principis que determinen el marc de gestió arxivística de la informació continguda en sistemes d'informació d'una organització.¹⁴⁷ I com a instrument de la política *Building trust in the public record (2021)* trobem la guia *Business System Assessment Framework (2018)*, basada en la part 3 de la norma ISO 16175, la qual detalla com establir un sistema de control i avaluació dels sistemes d'informació i aplicacions d'una organització. Aquesta guia proposa un procediment per identificar els sistemes d'informació, avaluar-los i incorporar-los a un inventari utilitzant un esquema que també proposa, el *Template for system information management plan (2018)*. Aquest model ha estat molt útil per donar forma al concepte de mapa de sistemes d'informació que presentem en la **GMAP**.

3.1.2.4 La dadificació de processos i el mapatge de dades

La idea d'una gestió documental basada exclusivament en dades és relativament nova en la literatura arxivística, tal com es detalla en l'apartat 2.3.6 de l'estat de la qüestió (capítol 2). Thomas Padilla afirmava fa pocs anys que "*to see collections as data begins with reframing all digital objects as data*" (Padilla, 2017), i Devon Mordell proposava que l'objectiu de l'arxivística ha de començar a ser "*renders the archive amenable to computational analysis*" (Mordell, 2022). A manca d'un major avenç doctrinal, no hi ha massa referents metodològics que connectin d'una forma sistemàtica el contingut de documents-objecte amb dades estructurades (Parilla et al., 2017), més enllà de la gestió de metadades i d'alguna aportació experimental (Makhlouf Shabou et al., 2020).

El model que hem utilitzat com a base de la **GDAD**, i de part de la conceptualització del model de dades, és el que formularen Antoni Puig-Pey i Pere Guiu a finals dels anys 90, i que encara constitueix la base de la formació d'arxivers en gestió de processos a bona part de les escoles d'arxivística espanyoles (Puig-Pey & Guiu, 2002, 2008). Al capítol IV del seu manual apuntaven el concepte de la graella de variables d'un procés, i detallaven una sistemàtica per identificar les dades dins dels documents d'un circuit documental i disposar-les en una matriu que en fes possible l'anàlisi i el posterior processament informàtic. Aquesta metodologia, que denominaren APS, s'adreça directament a la identificació de variables d'informació, no mers valors d'aquestes variables, amb

¹⁴⁶ *Information and data governance framework*, disponible a <https://www.naa.gov.au/about-us/who-we-are/accountability-and-reporting/information-and-data-governance-framework>.

¹⁴⁷ *Information Management Standard for Australian Government*, disponible a <https://www.naa.gov.au/information-management/standards/information-management-standard-australian-government>.

l'objectiu d'establir un patró que descrigui de forma estructurada un document que no necessàriament ha de ser estructurat. La **GDAD** correspon doncs a una evolució dels fonaments apuntats per la metodologia APS.¹⁴⁸

En el procés d'extracció de variables i caracterització de les dades s'ha vist en la necessitat d'incloure un element articulador d'aquestes dades que, en l'explotació posterior, s'ha revelat determinant per a la governança de la dada: l'objecte descrit. Aquest element forma part de l'estructura dels sistemes de governança de la dada, i es materialitza amb el que s'anomena "*business glossary*" (Gandomi & Haider, 2014; Loshin, 2015; Kononow, 2019). Però encara no s'ha consolidat una integració normalitzada entre els glosaris de negoci de les organitzacions i els models de descripció de conjunts de dades i ontologies més estesos (Chen 2022; Reina 2023).

Per a la identificació i classificació dels objectes descrits per les dades en aquestes dues guies hem utilitzat alguns conceptes extrets de l'ontologia PROV-O,¹⁴⁹ i s'ha revelat molt útil un plantejament no previst a la metodologia APS durant el buidatge de les variables dels documents: la identificació prèvia dels agents, les coses i les accions que poden aparèixer descrites en aquella evidència d'una forma discursiva, tal com es fa en la construcció de triplets RDF¹⁵⁰ i en el modelat dimensional de sistemes d'explotació de dades (Ur et al., 2015).

3.1.3 Llicències terminològiques

En ambdues guies s'utilitzen conceptes propis de diverses disciplines, els quals en cada disciplina tenen també sinònims no sempre acceptats per la bibliografia. En el model de dades del capítol 4 s'ha fixat la terminologia de referència per a cadascun dels conceptes, juntament amb la seva definició. Però això no obsta per a que, als efectes de garantir una bona comprensió de la guia, s'utilitzin en determinats casos sinònims o termes equivalents, de forma total o parcial. Es detallen a continuació algunes d'aquestes llicències terminològiques.

En el punt 2.2.1 de l'estat de la qüestió (capítol 2) es descriu la variabilitat que envolta el terme "document", en el nostre cas "document administratiu", i la difícil equivalència amb el terme anglès "*record*". Per evitar dubtes sobre l'abast del terme document en formes documentals no convencionals, i també en agrupacions o conjunts de dades

¹⁴⁸ Algunes propostes de Puig-Pey i Guiu, com la de considerar documents d'un circuit documental les pantalles o interfícies d'entrada i visualització de dades, les trobem també apuntades per altres autors (Yeo, 2012).

¹⁴⁹ PROV-O: *The PROV Ontology*, disponible a <https://www.w3.org/TR/prov-o/>.

¹⁵⁰ RDF 1.1 *Primer*, disponible a <https://www.w3.org/TR/rdf11-primer/>.

estructurades, vam optar per adoptar una solució similar a la del terme “referència documental” en la metodologia APS (Puig-Pey & Guiu, 2002, 2008), i vam decidir utilitzar el terme “evidència” (conceptualment “evidència documental”) per a designar qualsevol manera de deixar constància d’una acció o esdeveniment de forma persistent. El nostre terme “evidència” descriu tant un document tradicional, entès com un conjunt d’informació autònom i independent de qualsevol sistema que es pot gestionar com una unitat, com també qualsevol agregat de dades amb valor d’evidència d’un fet o acció. Però sovint en el text hem de fer la distinció entre aquestes dues formes documentals, per la qual cosa en alguns casos apareix el terme “evidència-objecte”, i també “document-objecte”, per indicar que l’evidència de la qual es parla és equivalent a un document tradicional.

Per afavorir la intel·ligibilitat del text tampoc s’ha eliminat el terme “document” de determinades construccions. Així doncs, per designar el flux d’evidències o documents dins un procés s’utilitza el terme “circuit documental” de forma indistinta amb el de “circuit evidencial”, utilitzant aquest segon especialment quan el circuit ja ha incorporat formes basades en dades estructurades.

El concepte de “sèrie documental”, malgrat pot estar formada per tot tipus d’evidències, i al final de la dadificació únicament per una matriu de variables, també conserva l’epítet “documental”, a fi d’afavorir-ne la comprensió i evitar sinonímies errònies amb la paraula “sèrie”, utilitzada en altres contextos.

El concepte de “procés” també és procliu a múltiples interpretacions, especialment tenint present que en la literatura arxivística no s’utilitza de forma habitual, i on enlloc de procés es parla de funcions, activitats, accions i altres nivells d’activitat. En aquest cas s’ha volgut unificar l’ús en un sol terme per reforçar el trencament amb estructures jeràrquiques i garantir la unicitat del concepte de procés al llarg del document.

Amb el concepte de “sistema d’informació” també s’ha cercat aquesta constància, llevat d’alguns casos on s’utilitza també el terme “aplicació”, per assenyalar especialment una manifestació concreta dels sistemes d’informació. En tot cas, en ambdues guies “aplicació” queda conceptualment inclosa dins el terme “sistema d’informació”.

Finalment, el concepte de “dada” és un dels més difícils de determinar terminològicament. L’abús del terme “dada” en la literatura recent el converteix en un terme polisèmic, tal com s’ha analitzat en els punts 2.2.1 i 2.2.2 de l’estat de la qüestió (capítol 2), i ha calgut concretar-lo amb termes més específics i amb algunes concessions. Així doncs, en el model de dades del capítol 4 s’utilitzen dos conceptes: la “dada elemental”, o unitat mínima d’informació que es pot gestionar i que pot prendre múltiples valors, i l’aparició d’aquesta dada en una evidència o sistema d’informació en concret, que és el que anomenem “ocurrència” de la dada. Però en el redactat de les guies s’incorporen algunes variacions. El terme “dada elemental” també se cita simplement com a “dada”, i quan apareix en evidències també s’anomena “variable d’informació”, en una concessió a la terminologia de la metodologia APS. En aquest

sentit, els termes “ocurrència” i “variable” es poden considerar sinònims. És important tenir present que el terme “dada” l'utilitzem sempre a un nivell abstracte, de mapa de dades, i que no és en cap cas sinònim de “valor” de la dada.

3.2 Guia d'inventari i mapatge (GMAP)

3.2.1 Estructura i mètode d'aquesta guia

Aquesta guia està directament vinculada als conceptes i elements del model de dades detallat al capítol 4. Per aquest motiu, s'indica en primer lloc, de forma gràfica, el pas a pas de la metodologia per a la construcció del model, i en el mateix gràfic els vincles amb cadascun dels elements (entitats o relacions) del model de dades.

A continuació es detalla el desenvolupament de cadascun dels passos de la metodologia. Per a cada pas, es descriu:

- a) La definició de l'entitat que és objecte de mapatge.
- b) El raonament a seguir per identificar-la de forma clara i unívoca.
- c) El procediment per a caracteritzar-la a través de la identificació i descripció dels seus atributs principals.
- d) El procediment per establir les relacions amb altres instàncies de la mateixa entitat, i també amb altres entitats del model.
- e) El raonament a seguir quan es creïn noves entitats, es modifiquin o causin baixa del model.

Els indicadors i elements de mesura de la qualitat del mapa es descriuen en el capítol 5. Les operacions que poden ajudar a millorar la qualitat de les entitats identificades o de les relacions amb altres instàncies de la mateixa o d'altres entitats, un dels potencials d'aquesta metodologia, es comenten en les conclusions d'aquesta tesi (capítol 8).

L'element determinant d'un bon mapatge és l'exhaustivitat en la recollida de la informació i el rigor en el seu buidatge i anàlisi. Per aquest motiu, s'inclou en el punt 3.4 d'aquest capítol una proposta procedimental per organitzar i sistematitzar la recollida d'informació prèvia a cada acció de mapatge, i especialment a la que té més risc de subjectivitat: la identificació dels processos i la construcció del mapa de processos.

3.2.2 Model conceptual per a la construcció dels mapes

La construcció dels diferents mapes que configuren el mapa global es pot fer de dues maneres:

- A. **Enfocament *top-down***. Consisteix en identificar les diferents entitats a partir d'un treball de camp general d'inventari i catalogació, i a continuació construir el mapa de relacions i preparar la base per un detall posterior.
- B. **Enfocament *bottom-up***. Consisteix en treballar a nivell de procés, sèrie o sistema, identificant totes les entitats implicades, i desgranar-les fins al màxim nivell de detall. Aquest treball, repetit per cadascuna de les entitats, anirà alimentant el mapa global de l'organització.

Els moments per a construir aquests mapes, combinant ambdós enfocaments, poden ser tres:

- A través de **l'anàlisi de sistemes**, quan es desenvolupen aplicacions informàtiques. El desenvolupament d'una solució tecnològica implica habitualment una lectura funcional prèvia, la definició de requeriments funcionals, i un disseny conceptual que després es tradueix en una implementació física/tecnològica concreta. Aquest disseny conceptual analitza les activitats que es porten a terme, en quina seqüència ho fan, quines evidències i/o dades generen, i d'altres requisits relacionats amb l'accés i la conservació d'aquesta informació. El desenvolupament d'aplicacions ofereix, per tant, la primera porta d'entrada a la confecció d'aquests mapes.
- A través de **l'anàlisi dels processos**, quan es fa reenginyeria de processos. La reenginyeria de processos, sovint incardinada dins la implantació de sistemes de gestió de la qualitat, parteix habitualment d'una identificació sistemàtica dels processos, el mapa de processos, per aprofundir després en el detall de cada procés. En aquest camí analitza les tasques en les quals es desglossa cada procés, les evidències i dades que generen aquestes tasques, en quins sistemes d'informació es recullen, i quin paper hi juguen els diferents agents del procés. La reenginyeria de processos ofereix, doncs, una segona porta per entrar a la construcció dels mapes que comentem.
- A través de **l'anàlisi dels documents**, quan s'implanta un sistema de gestió documental. Tot i ser la tasca menys habitual, cada cop és més freqüent que la implantació de sistemes de gestió documental parteixi d'una identificació sistemàtica de les sèries documentals que es generen, sovint identificades amb els processos, i d'una descripció de les característiques formals i funcionals de l'expedient tipus de cada sèrie. Això inclou les evidències que el componen, les

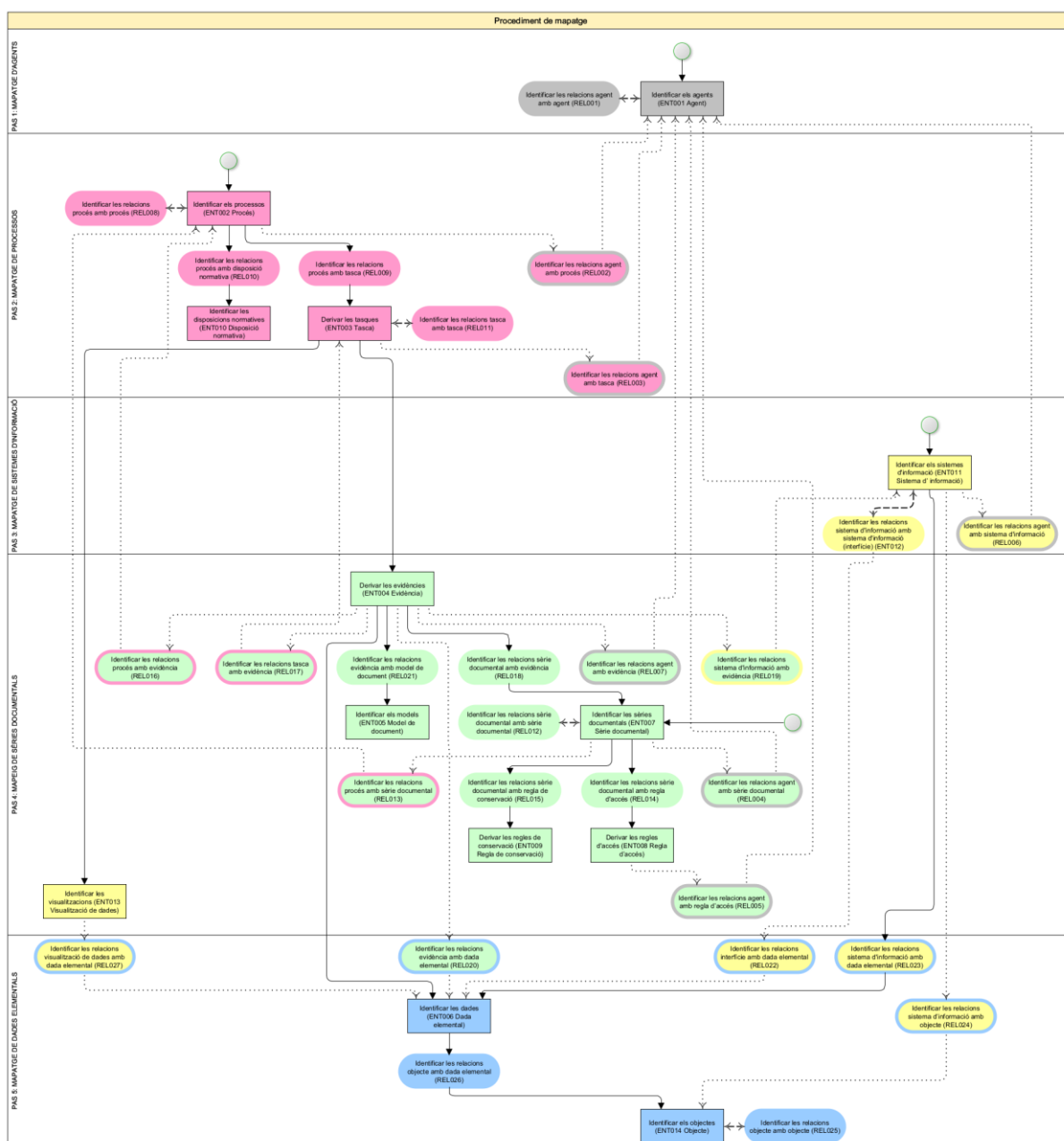
dades i metadades que l'acompanyen, els sistemes d'informació en els quals ha de residir o amb els quals interacciona, els requisits d'accés i ús, i els períodes durant els quals s'ha de conservar l'expedient. Tot això ofereix per tant un tercer escenari per a la construcció dels mapes que es descriuen a continuació.

Habitualment es construirà una primera versió del mapa global a alt nivell, i es baixarà fins al nivell de detall que el projecte inicial permeti. I a partir d'aquí, s'utilitzaran les oportunitats organitzatives que s'ha comentat per anar completant, detallant i actualitzant cadascun dels mapes que componen el mapa global. És per aquest motiu que aquesta guia es planteja per a ser aplicada preferentment des d'una aproximació *top-down*, i des d'una perspectiva neutra pel que fa al moment i oportunitats de construcció dels mapes. Per conèixer l'aproximació *bottom-up*, orientada a la transformació, vegeu la **GDAD**. Ambdós models es combinen en un mateix entorn de governança, i es complementen en l'assoliment de transformacions en el substrat informacional de l'organització.

Es presenta a continuació un resum estructurat del conjunt de la metodologia que es proposa, incloent les referències als elements del model de dades que s'informen en cada fase. En aquest diagrama es representen, per cadascuna de les fases de la metodologia, les entitats que s'identifiquen i les relacions que s'estableixen entre elles. En alguns casos, les entitats són producte de la descomposició d'entitats prèvies, i en aquest cas s'utilitza el terme "derivació".

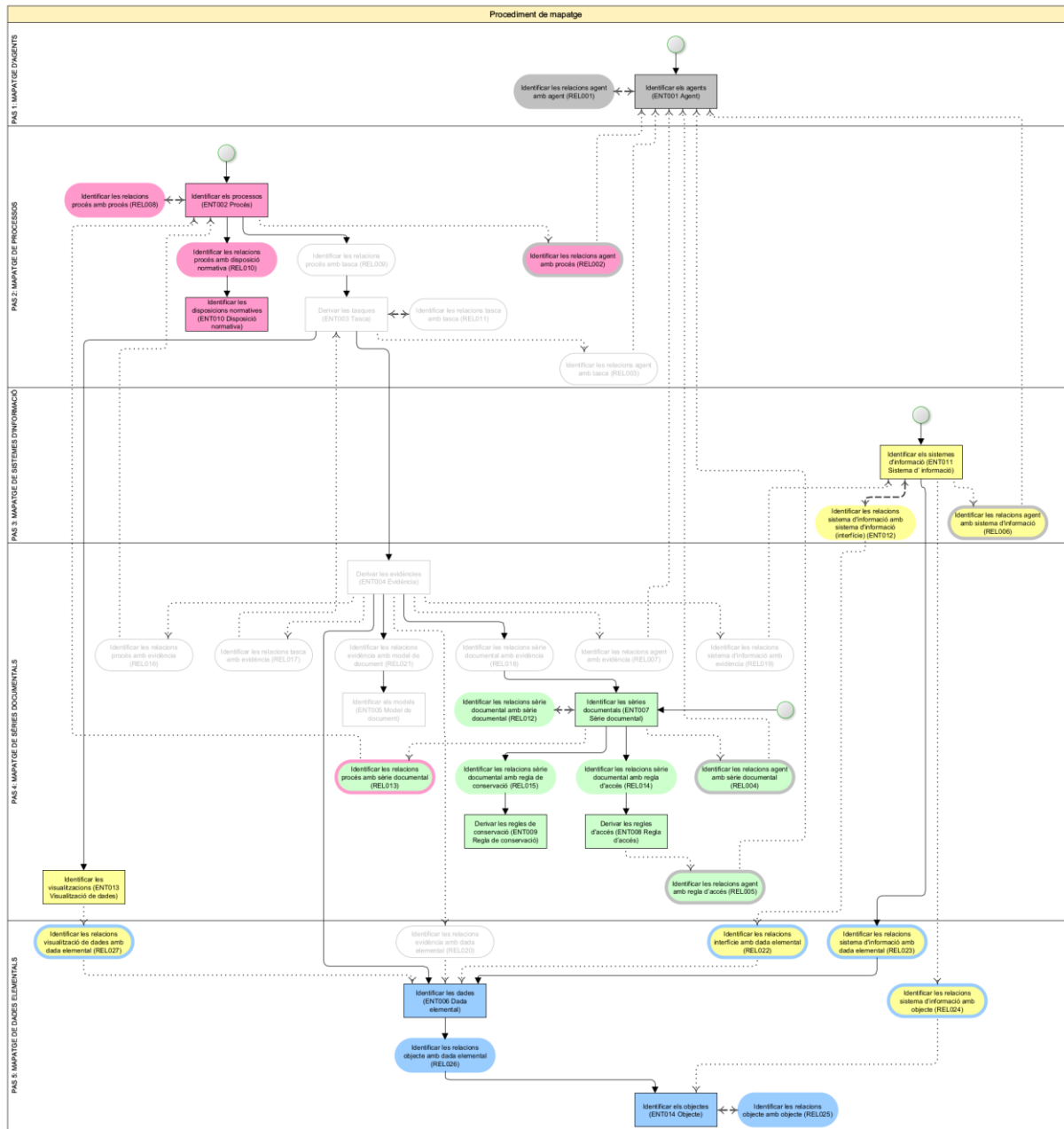
A través de colors s'indiquen també les famílies d'entitats, que s'agrupen en cada cas al voltant d'un domini o entitat principal, indicant en la vora de les relacions els vincles amb altres dominis.

Figura 3-1: Estructura general de la metodologia



Com hem comentat en la introducció, aquesta metodologia contempla dues possibles aproximacions a la realitat que es vol descriure i governar. La primera, objecte d'aquesta guia, és l'aproximació *top-down*, que parteix de la identificació general de les diferents entitats a partir d'un treball de camp, i a partir d'aquí les categoritza i n'identifica les relacions. La part de la metodologia que correspon a aquesta aproximació és la que s'indica en el següent diagrama:

Figura 3-2: Estructura de la Guia d'inventari i mapatge (GMAP)



Descrivim a continuació el detall de cada pas de la metodologia.

3.2.3 Pas 1: mapatge dels agents¹⁵¹

Figura 3-3: Estructura per al mapatge dels agents



3.2.3.1 Què s'entén per agent

Definirem agent com aquell ens que pot intervenir activament o passiva en un flux d'informació. S'entén per participació activa la creació i tramesa d'informació, i per participació passiva la recepció i arxiu de la informació.

Els agents poden ser humans o màquines. Els agents humans poden ser persones físiques (individus), persones jurídiques (conjunts d'individus amb existència simbòlica comuna) o bé rols (persones físiques o jurídiques tipus, identificades a partir de l'abstracció de les seves característiques comunes).

Quan els agents són màquines, poden tenir igualment un paper actiu o passiu. Actualment hi ha un debat important sobre si els agents no humans, i per tant no responsables de les seves activitats, han de ser considerats "agents" d'un procés o no (vegeu punt 4.6.1.1.1). En aquesta metodologia, i davant la manca d'un consens bibliogràfic unívoc, s'opta per no considerar com a agents els sistemes d'informació i aplicacions. De fet, en aquesta metodologia només s'inclou la faceta passiva dels sistemes d'informació, aquella derivada del fet que emmagatzemen informació. Els agents actius, siguin aplicacions semi-autònomes (amb interfície d'usuari), aplicacions completament autònomes (RPA) o sistemes intel·ligents (basats en *machine learning* o intel·ligència artificial), no es consideren agents, sinó instruments (eines) que actuen per delegació d'agents humans. El motiu rau en el fet que la consideració d'agent ve determinada pel compliment de dos requisits sobre l'acció o tasca realitzada:

- Tenir-ne la competència o qualsevol altre tipus d'encàrrec formal no ocasional
- Esdevenir responsable del resultat o producte generat

¹⁵¹ La construcció dels mapes sempre es fa des d'una perspectiva concreta, la de l'organització on s'implanta el sistema de governança de la informació. No s'utilitza una visió neutra, sinó sempre una visió de part, cosa la qual implica que el detall en els elements interns (agents, evidències, etc.) sempre serà molt superior al detall i concreció en els elements externs.

Atès que la competència i la responsabilitat només poden ser exercides i assumides per agents humans, les tasques portades a terme per agents no humans de forma automàtica o desatesa s'imputaran a l'agent humà que hagi estat el promotor¹⁵² del desenvolupament d'aquell sistema o aplicació, o bé que sigui el responsable o ens competent en l'execució d'aquella tasca.

Quan fem referència a agrupacions documentals arxivístiques, l'agent també rep el nom de "agent productor", o simplement de "productor". En el marc d'aquesta metodologia, utilitzarem agent com a productor només quan el relacionem amb sèries documentals.

3.2.3.2 Com s'identifica un agent

El concepte d'agent és un concepte abstracte. En el marc d'aquesta metodologia no s'identifiquen persones o usuaris concrets, sinó persones jurídiques o rols (càrrecs, perfils singulars, etc.). Aquestes persones jurídiques o rols estaran instanciades en el món real en persones físiques, però mai s'identificarà com agent en el mapa a una persona física concreta.

Hem comentat abans la necessitat de tenir en compte que el mapatge es fa des d'una perspectiva concreta, la de l'organització on s'implantarà el sistema de governança de la informació. Des d'aquesta perspectiva, doncs, cal distingir entre agents interns (o pertanyents a l'organització) i agents externs (o aquells amb qui l'organització interacciona).

La identificació dels agents interns es fa habitualment a partir dels documents que assignen algun tipus de competència o responsabilitat sobre una acció o tasca d'una organització. Aquests documents poden ser normativa (interna o externa), manuals de procediment, instruccions o encàrrecs escrits. En tot cas, tot agent ha de tenir existència documentada en l'organització, sigui en forma de normativa, sigui en forma de presència recurrent d'un determinat rol en procediments. Sense aquesta existència documentada no es podran incorporar agents al mapa. No s'hi faran constar, per tant, elements consuetudinaris sense base escrita (normativa o equivalent). En cas que es tingui coneixement de tasques desenvolupades per persones no assimilables a agent (sense base escrita de la seva activitat), aquelles tasques s'assignaran a l'organització en general, i es considerarà tota l'organització com agent (agent corporatiu).

La identificació dels agents externs requereix conèixer prèviament els processos de l'organització. Els agents externs poden ser:

- Proveïdors d'un producte, servei, informació, etc. de l'organització.

¹⁵² El concepte de promotor d'un sistema o aplicació és diferent del concepte de desenvolupador, i s'explica en el punt corresponent a la caracterització de sistemes d'informació.

- Clients d'un producte, servei, informació, etc. de l'organització.
- Supervisors d'un producte, servei o activitat de l'organització.

Per tant, serà a partir de les entrades i sortides dels processos que es podran identificar els agents externs. Per tant:

- Els **agents interns** s'inclouran en el mapa abans de portar a terme la identificació dels processos, o en tot cas de forma independent a l'elaboració del mapa de processos.
- Els **agents externs** s'inclouran en el mapa com a resultat de la caracterització i mapatge dels processos.

Un cop identificat, l'agent generarà una instància de l'entitat **ENT001 Agent** del model de dades.

3.2.3.3 Com es caracteritzen els agents

La caracterització dels agents es fa a partir de les fonts abans esmentades. Els atributs que caldrà informar per a cada agent en el moment de la seva identificació són els següents:

- **ENT001-001 Identificador**
- **ENT001-002 Nom**
- **ENT001-003 Tipus**
- **ENT001-004 Contacte**
- **ENT001-005 Observacions**
- **ENT001-006 Data creació**

Els criteris per informar cada atribut es troben descrits en el model de dades (capítol 4).

3.2.3.4 Com es relacionen els agents entre ells

Els agents es relacionen entre ells d'acord amb la relació **REL001 agent amb agent**. Aquesta és una relació amb atributs, que pot prendre la següent significació:

- **Agent** Pot dependre/ser staff de un o més **agents**

Aquesta relació només es desenvolupa per als **agents interns** de l'organització, i excepcionalment s'hi poden incloure agents externs, sempre que tinguin una relació jeràrquica amb un agent intern.

Els agents d'una organització poden tenir entre ells els següents tipus de relació:

1. **Relació de dependència jeràrquica.** Un agent pot dependre jeràrquicament d'un altre agent. La jerarquia condiona la capacitat per a prendre decisions, i l'àmbit funcional en què es pot exercir aquesta capacitat. Habitualment les organitzacions són monojeràrquiques, de manera que la relació de dependència jeràrquica només té un element pare. Això vol dir que un agent pot tenir un agent de qui depèn (valor "Depèn de" de l'atribut **REL001-001 Rol de l'agent**) o un agent que en depèn (valor "És staff de" de l'atribut **REL001-001 Rol de l'agent**).
2. **Relació de col·laboració.** Un agent pot col·laborar de forma habitual amb d'altres agents. Aquesta relació habitualment es materialitza en la participació en uns mateixos processos. En conseqüència, serà un tipus de relació que no es reflectirà en el mapa d'agents, i que no s'ha contemplat com a valor en l'atribut **REL001-001 Rol de l'agent**.

El nivell superior d'agent, en el cas de les relacions de dependència jeràrquica, és el conjunt de l'organització.

Per a la consistència del mapa i les posteriors relacions amb altres entitats del model, és necessari recollir tots els agents que conformen o interactuen amb l'organització. Però als efectes de representació del mapa d'agents de l'organització, habitualment només s'inclouran en el mapa els agents interns.

3.2.3.5 Com es relacionen els agents amb altres entitats

En aquest pas de la metodologia només es contempla la necessitat d'informar la relació **REL001 agent amb agent**.

La resta de relacions de l'entitat agent es van completant en els passos següents de la metodologia:

- **REL002 Agent** pot ser responsable/destinatari d'un o més **processos**
- **REL003 Agent** pot ser executor d'una o més **tasques**
- **REL004 Agent** pot ser productor d'una o més **sèries documentals**
- **REL005 Agent** pot obtenir permisos d'accés d'una o més **regles d'accés**

- **REL006 Agent** pot ser promotor d'un o més **sistemes d'informació**
- **REL007 Agent** pot ser creador/receptor d'una o més **evidències**

3.2.3.6 Com es gestionen les altes, baixes i modificacions d'agents

L'estructura d'agents d'una organització va evolucionant amb el temps. Per tant, la definició dels agents està subjecta a temporalitat. El mapa d'agents ha d'incloure els agents que ja no són actius, però que van existir durant un temps vinculats a processos i documentació generada.

La casuística de cicle de vida relacionada amb els agents és la següent:

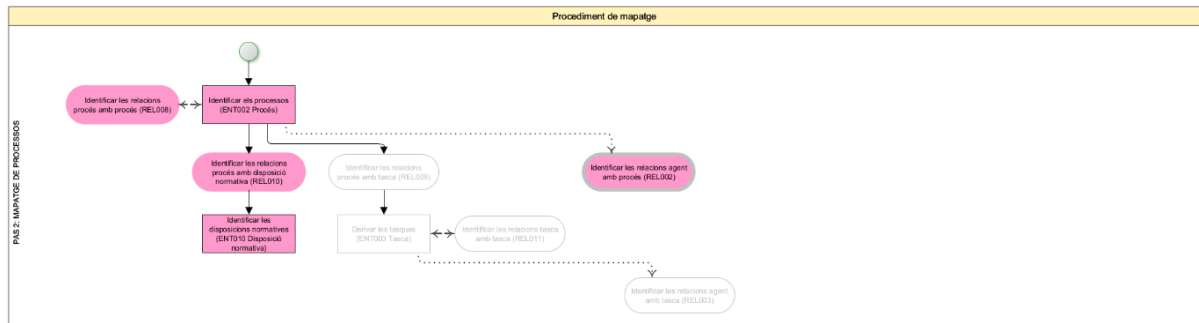
Taula 3-1: Pautes d'actuació per al manteniment del mapa d'agents

ESDEVENIMENT	ACTUACIÓ
Desaparició d'un agent	S'informarà l'atribut ENT001-007 Data supressió amb la data en que l'agent deixa de tenir existència jurídica i/o funcional.
Absorció d'un agent	Es donarà de baixa l'agent absorbit com s'ha indicat en el punt anterior, i es donarà d'alta el nou agent, indicant en ambdós casos la relació a l'atribut ENT001-005 Observacions .
Canvi de nom d'un agent	Cada organització decidirà, en funció de la rellevància del canvi, si únicament modificarà l'atribut ENT001-002 Nom , o bé donarà de baixa l'agent i es tornarà a donar d'alta amb el nom nou i el canvi informat a l'atribut ENT001-005 Observacions .
Canvi de funcions d'un agent	Un canvi de funcions no provocarà mai la baixa d'un agent, atès que quedarà reflectida pel seu vincle amb determinats processos (relació RELO02), que sí que es modificarà.

ESDEVENIMENT	ACTUACIÓ
Canvi de dependència d'una agent	Un canvi de dependència d'un agent es materialitzarà amb una baixa de la relació vigent REL001 agent amb agent , i amb l'alta d'una nova relació que reflecteixi la nova dependència.
Canvi de naturalesa jurídica d'un agent	Cada organització decidirà, en funció de la rellevància del canvi, si únicament modificarà l'atribut ENT001-003 Tipus , o bé donarà de baixa l'agent i es tornarà a donar d'alta amb el nom nou i el canvi informat a l'atribut ENT001-005 Observacions .
Canvi de persona/adreça de contacte d'un agent	Un canvi de persona/adreça de contacte d'un agent només provocarà la modificació de l'atribut ENT001-004 Contacte , sense cap altre efecte.

3.2.4 Pas 2: mapatge dels processos

Figura 3-4: Estructura per al mapatge dels processos



3.2.4.1 Què s'entén per procés

Definirem procés com el conjunt d'activitats d'un o diversos agents adreçades a aconseguir un objectiu comú, generades per una entrada (*input*) d'un proveïdor, i orientades a la producció d'un producte o servei (*output*) cap a un client (intern o extern).

A l'estat de la qüestió (apartat 2.3.2 del capítol 2) es detalla el debat al voltant del concepte de procés en l'àmbit arxivístic i de la gestió documental.

3.2.4.2 Com s'identifica un procés

La representació d'un procés és una visió limitada d'una realitat funcional global, que ve determinada per on se situa l'acció d'inici i l'acció de fi del procés. Aquesta ubicació és sovint convencional, i reflecteix la visió funcional de qui analitza i representa el procés. Aquesta visió és a la seva vegada determinada per la concepció que tingui l'analista de com s'ha d'estructurar funcionalment una organització, i en últim terme d'aspectes de cultura i de valors de l'organització. Per aquest motiu, és especialment important intentar objectivar al màxim com s'identifica un procés.

La identificació dels processos es pot fer a través de tres tècniques:

1. Recollir com a procés allò que la pròpia organització diu que és un procés (aquell conjunt d'activitats al qual l'organització li dona un nom concret).
2. Recollir com a procés allò que cànons externs marquen com a procés, per exemple la doctrina sobre com funciona una empresa o una Administració pública.

3. Recollir com a procés allò que va d'una activitat que no té antecedent objectiu, fins a una activitat que no té conseqüent objectiu.

El tercer sistema és sens dubte el que sembla més objectiu, però depèn del rigor amb què es defineixi cada activitat. Per tant, què pot esdevenir una activitat d'inici de procés?

1. Aquella activitat que s'inicia com a resultat d'una acció voluntària d'un agent (per exemple una petició o una comanda).
2. Aquella activitat que s'inicia com a resultat d'una detecció d'un fenomen extern (per exemple un sensor de moviment o una programació de calendari).

Un requisit de l'activitat d'inici és que genera una cadena de causa-efecte, és a dir que provoca una activitat conseqüent, sigui per una obligació normativa d'actuar, sigui perquè dispara qualsevol tipus d'automatisme.¹⁵³

Pel que fa a l'activitat de fi, es pot identificar:

1. Perquè és una activitat que no genera cap efecte, cap obligació de seguir actuant.
2. Perquè és una activitat que afecta l'agent que ha iniciat el procés (en el cas dels processos iniciats com a resultat d'una acció voluntària).

Per eliminar doncs aquest risc de subjectivitat, en aquesta guia es proposa utilitzar un mètode habitualment utilitzat per a l'elaboració de quadres de classificació funcionals en sistemes de gestió documental, però amb algunes especificitats.¹⁵⁴ En el nostre model hem adoptat un enfocament relacional, on no existeixen com a tals relacions d'inclusió, i les relacions que s'estableixen són de caràcter estrictament funcional (consecutivitat, complementarietat, etc.). Per sobre del nivell de procés tan sols utilitzem una categoria, la que correspon al concepte "família" de processos, i només a títol orientatiu, atès que és una categoria molt poc consistent, mentre que per sota del nivell de procés ja ubiquem el nivell més operatiu, el de tasca. Això no obsta per poder utilitzar com a pauta de treball un mètode que ha demostrat la seva eficàcia en l'elaboració de nombrosos models funcionals, i que està reconegut com un estàndard internacional (ISO 26122).

Segons la norma ISO/CD 26122, hi ha dos tipus d'anàlisi per a la identificació de funcions i activitats en una organització. El primer és l'anàlisi funcional (o jeràrquic), que va del que és general al que és específic, i que consisteix a descompondre en parts i subparts lògiques les funcions d'alt nivell de l'organització. El segon tipus és l'anàlisi

¹⁵³ La derivada d'aquesta concepció és que es podria arribar a entendre l'organització com un únic procés continuu, en el moment en què anem relacionant tots els processos entre ells sense solució de continuïtat. La fragmentació en processos es considera doncs convencional.

¹⁵⁴ El procés habitual d'identificació de processos es desplega de forma jeràrquica, del més general al més específic, desenvolupant-se per nivells de concreció i "retallant" els processos segons unes dimensions homogènies. Així doncs, es parla de macroprocessos, processos, subprocessos, etc.

seqüencial (o de processos), que parteix de l'anàlisi en detall de com l'organització desenvolupa les seves activitats, per posar progressivament en relació aquestes activitats fins a construir el mapa de processos o funcional. En el nostre model combinem ambdós tipus d'anàlisi: l'anàlisi funcional, a tres nivells, per a la construcció inicial del mapa de processos, i l'anàlisi seqüencial, per a l'ajustament, interrelació dels processos i posterior manteniment del mapa.

L'anàlisi funcional es basa en aplicar aquest patró de descomposició:

Figura 3-5: Patró de descomposició de l'anàlisi funcional



El mètode que proposem aplicar consta de cinc operacions:

- Buidatge de funcions i activitats
- Identificació i eliminació de duplicitats
- Establiment de relacions
- Validació mitjançant iteracions
- Codificació i denominació definitiva

Es descriu a continuació el detall de cadascuna de les operacions.

3.2.4.2.1 Buidatge de funcions i activitats (processos)

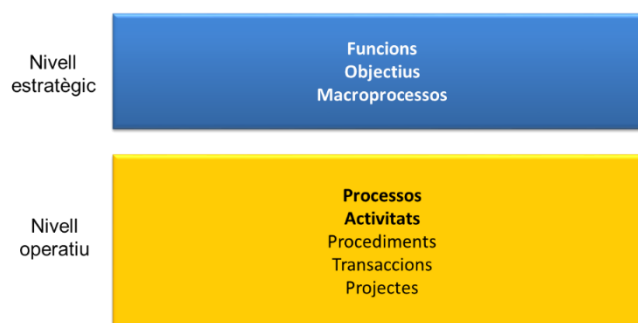
La fase de l'elaboració d'un sistema de classificació funcional on es posa a prova el criteri de l'analista és la detecció de les funcions i activitats. Un analista experimentat podrà intuir l'existència de determinades funcions i activitats tan sols amb el coneixement del tipus d'organització, i podrà aplicar patrons de descomposició de funcions i relacions entre activitats apreses a partir de l'anàlisi d'organitzacions de característiques similars. Tanmateix, fins i tot en aquells casos on més clar sembli quines són les funcions i activitats, és important mantenir una sistemàtica en el desenvolupament del procés d'anàlisi per garantir-ne el rigor i evitar apreciacions subjectives.

Per aquest motiu és important tenir molt clar què s'entén per funcions i activitats, i com es presenten sovint a les fonts. De forma general podem entendre per:

- **Funció:** qualsevol finalitat, responsabilitat o tasca d'alt nivell assignada a una entitat col·lectiva per la legislació, política o manament amb la finalitat de retre comptes. La funció és la unitat més gran d'activitat de qualsevol organització. En el nostre cas, l'assimilarem a família de processos.
- **Activitat:** tasca duta a terme per un organisme per fer una de les seves funcions. Les activitats poden estar associades a una funció o a diverses. En el nostre cas, tot i mantenir la terminologia ISO 26122, l'assimilarem al concepte de procés.
- **Transacció:** la unitat més petita d'activitat d'una organització. Cada transacció dona lloc a la producció d'un o més documents d'arxiu. És important destacar que les transaccions sempre són accions, no objectes de l'acció ni tipus de documents. En el nostre cas utilitzarem per aquest nivell el concepte de tasca.

En aquesta anàlisi ens ajudarà establir un patró de sinonímia en dos nivells:

Figura 3-6: Nivells inicials per a la identificació de processos



Per aplicar de manera més àgil aquest patró de sinonímia, és útil fer la següent operació. Es tracta de preparar un quadre de buidatge amb cinc columnes, una per a cada concepte sinònim, ordenats de més genèric a més específic, de la manera que s'indica:

Figura 3-7: Estructura per al buidatge dels processos



A mesura que es buida cada font d'informació, es van anotant a la columna corresponent les activitats o processos detectats, citant per a cada activitat la font de la qual s'ha obtingut mitjançant la seva referència al registre de fonts. L'ordre de buidatge de les fonts, tal com s'ha establert en la seva prioritització, facilita que la informació s'obtingui a la seqüència que s'indica:

- En primer lloc, els objectius principals de l'organització i les seves funcions bàsiques, allò que fa única l'organització.
- En segon lloc, el desenvolupament complet de les funcions de l'organització. Tot i les aparences, les funcions generalment no es corresponen amb l'estructura orgànica de l'organització, són molt més estables i es veuen poc afectades per reestructuracions. Tanmateix, és convenient testejar amb l'estructura orgànica la detecció de totes les funcions principals.
- En tercer lloc, la descomposició de cada funció en diferents activitats o processos. És important que les activitats tinguin límits clars, i que sigui possible determinar quina acció inicia una activitat, i quina la finalitza. Les activitats no s'han de definir a partir de la manera com s'organitza la documentació de l'organització (per exemple, deduint que hi ha una activitat de "resolució" pel fet que trobem expedients que agrupen cronològicament diferents tipus de resolucions administratives); aquesta organització és circumstancial, en aquesta fase res no ens assegura que sigui correcta o lògica, i probablement podrà ser modificada com a resultat del redisseny dels processos.
- I finalment la descomposició de cada activitat en diferents transaccions o procediments. Aplicarem en aquest cas els mateixos raonaments que al punt anterior.

Aquest enfocament estrictament deductiu no sempre és possible aplicar-lo. Res no assegura que sigui possible detectar totes les funcions abans de començar a identificar activitats, i fins i tot algunes funcions no es podran detectar més que a partir de la identificació prèvia de determinades activitats o transaccions. Per aquest motiu, proposem un model de treball iteratiu, basat en una anàlisi eminentment lingüística de les fonts escrites i en la contrastació sistemàtica amb les fonts personals.

Les mencions a funcions i activitats poden prendre a les fonts formes molt diverses. Algunes recomanacions poden ajudar a detectar amb més encert aquestes activitats:

- Les funcions i activitats es defineixen mitjançant verbs. El verb és, doncs, element bàsic per identificar el tipus d'activitat de què es tracta. És important no caure en el parany de fixar-se en els substantius. Per exemple: un organisme públic es pot dedicar a les arts escèniques. Tot i això, mai considerarem "arts escèniques" com una funció (almenys en aquesta fase de l'anàlisi). La promoció de les arts escèniques amateur mitjançant subvencions no té res en comú amb l'ordenació de les arts escèniques mitjançant la concessió d'autoritzacions a companyies de teatre amateur, i en canvi sí que té molt en comú amb la promoció de les arts plàstiques mitjançant subvencions. Es tracta de dues activitats totalment diferents, malgrat que l'objecte sobre el qual s'apliquen sigui el mateix.
- Sovint a les fonts documentals les activitats es descriuen en una forma substantivada. Així, en els objectius d'una organització serà més freqüent trobar

“promoció de les arts escèniques” que no pas “promocionar les arts escèniques”, i a la memòria es parlarà de “subvencions concedides a companyies teatrals” abans de “subvencionar companyies teatrals”. És important estar atent a aquests esments dissimulats d'activitats.

- Determinades funcions d'alt nivell no apareixen a les fonts de forma directa, però poden ser assignades a qualsevol organització. Serien, per exemple, el respecte i la protecció del medi ambient, etc. La decisió d'incloure-les o no al model funcional depèn del seu impacte en forma d'activitats sobre l'organització.
- En recollir les funcions i activitats que es vagin detectant, és convenient utilitzar inicialment un redactat prou explícit per evitar confusions. Per exemple, si es fa referència a una convocatòria de subvencions per conveni a companyies de teatre de carrer, seria incomplet recollir aquesta activitat simplement com a “subvencions” o “subvencions a companyies teatrals”. El més recomanable i clar, en aquesta fase del procés d'anàlisi, és citar l'activitat de manera completa i en infinitiu: “subvencionar mitjançant conveni a companyies de teatre de carrer”.
- Paral·lelament al quadre de buidatge, és convenient començar a elaborar la fitxa descriptiva de cada activitat. En aquesta fase, el model encara és molt inestable per invertir massa temps en la confecció d'unes descripcions detallades, però és convenient associar a aquelles entrades que sigui possible una nota explicativa que contingui informació sobre l'abast de la funció, activitat o transacció, els seus límits i contingut, els departaments o actors que la duen a terme, i la data i la norma d'inici.

3.2.4.2 Identificació i eliminació de duplicitats

Amb les operacions anteriors s'ha aconseguit elaborar un primer instrument d'anàlisi que anomenarem “model funcional”. Aquest model ens permetrà una sèrie d'operacions de depuració de la informació recollida, que de manera coherent ens ajudaran a construir el nucli del mapa de processos.

Per fer referència a les funcions i activitats detectades, contrastar duplicitats i registrar-ne l'eliminació, i posteriorment establir relacions entre aquestes funcions i activitats, cal assignar a cada entrada un codi o referència provisional. Una simple numeració seqüencial per cadascuna de les columnes emplenades serà suficient.

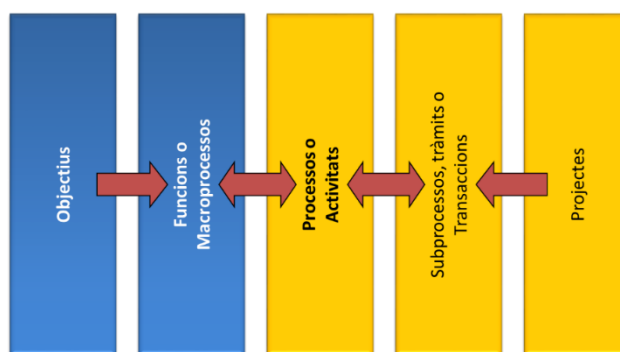
La primera operació per depurar la informació serà eliminar els duplicats. Els duplicats no apareixeran mai de forma directa, sinó en forma de sinònims. Així, per exemple, d'unes fonts generals podem haver obtingut l'activitat concertar convenis amb les companyies de teatre emergents, i també l'activitat donar ajudes econòmiques a les companyies novells de teatre. Un primer tipus de duplicitat pot ser: és el teatre emergent sinònim del teatre novell? Però també ens podem plantejar la duplicitat següent: Es

concerten convenis amb companyies teatrals emergents que NO tinguin per objectiu donar ajudes econòmiques?

Ho comprovarem amb un buidatge més detallat de les fonts, amb l'ampliació mitjançant fonts externes, amb la consulta de documents i expedients d'exemple, i finalment preguntant-ho a les persones responsables de la tramitació esmentada. És probable que aquest treball més detallat sobre aquestes dues activitats aparentment diferents ens porti a descobrir que totes dues consisteixen en el mateix, i fins i tot a consensuar amb els tramitadors una denominació més satisfactòria i clara, com ara "subvencionar mitjançant conveni a companyies de teatre amateur".

Aquesta eliminació de duplicats inclou també verificar que no hi ha solapament de contingut entre diverses entrades del mateix nivell. Les notes d'abast, que hem elaborat juntament amb el quadre de buidatge, ens ajudaran a detectar aquests solapaments i plantejar les possibles solucions.

Figura 3-8: Detecció de duplicitats en l'anàlisi funcional



El contrast amb les fonts d'informació és preceptiu abans d'eliminar els duplicats. Basar-se únicament en sinònimes aparents és arriscat, i pot implicar fer desaparèixer del model funcional l'únic rastre d'una activitat difícil de detectar. És recomanable actuar amb prudència durant aquesta fase i conservar la traça de les activitats eliminades fins a disposar de la seguretat completa de la duplicat.

3.2.4.2.3 Establiment de relacions

L'establiment de relacions entre les diferents funcions i activitats és l'element clau per a que el sistema de classificació esdevingui realment un mapa de processos. La potència del sistema de classificació la proporciona el fet que l'assignació d'un document o agrupació a una determinada classe comporta la seva contextualització dins d'un marc global d'activitats relacionades i interdependents. Aquestes relacions s'utilitzaran posteriorment tant en el procés de descripció com durant totes les operacions relacionades amb el cicle de vida de les entitats documentals.

La construcció d'aquestes relacions és una operació que mobilitza tota la informació recopilada fins ara, i que, de la mateixa manera que l'eliminació de duplicats, implica contrastar amb les fonts qualsevol modificació que apliqui l'analista sobre el model funcional.

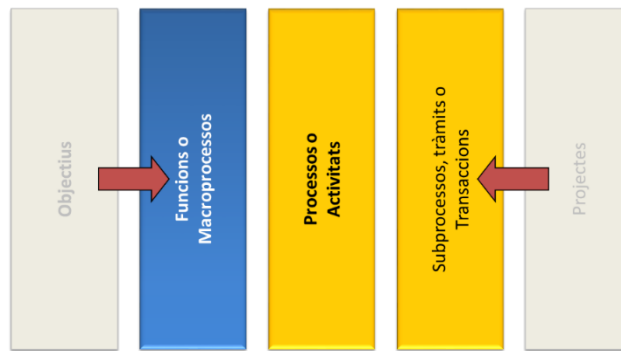
El primer pas que cal fer consisteix en eliminar la primera i última columnes del quadre de treball. Això no implica que el producte final tingui tres nivells, però garanteix que els nivells de treball nuclears corresponen a conceptes realment funcionals. De fet, el nombre final de nivells dependrà de la quantitat i la complexitat de les funcions exercides per l'organització, i es definirà amb l'objectiu de crear un mapa de processos còmode i comprensible per a tots els membres de l'organització.

Per tant, eliminarem:

- El primer nivell, corresponent als **objectius**, els quals ens poden indicar una correspondència amb les funcions o macroprocessos. Però sovint els objectius corporatius se centren només en els processos clau (obviant aquells processos d'execució obligada o rutinària), o bé són un desideràtum (sobretot els de caràcter més polític) que no es poden vincular a activitats concretes. L'analista haurà de buscar, per a cada objectiu, l'existència de funcions relacionades, i en el cas que no existeixin com a funcions però sí com a activitats, crear les funcions o macroprocessos corresponents.¹⁵⁵
- El darrer nivell, corresponents als **projectes**. En aquest cas es tracta d'activitats concretes, però realitzades de forma puntual. La inclusió dels projectes en el model funcional ens permet detectar activitats recurrents ocultes darrere d'una gestió individualitzada o per projectes. En el mapa de processos només s'han d'incloure activitats que es realitzin de forma recurrent o regular, no activitats puntuals. L'anàlisi de les característiques dels projectes permetrà detectar aquestes possibles regularitats, i a partir d'aquí l'existència de potencials processos. L'analista suprimirà aquesta columna quan tots els projectes s'hagin incorporat com a processos o activitats, o quedin potencialment inclosos en activitats ja detectades com a concrecions o desenvolupaments puntuals d'aquestes.

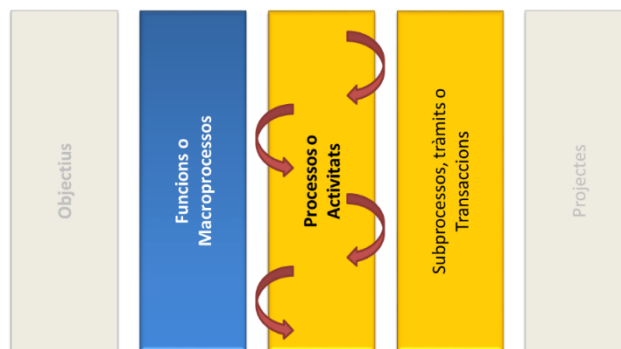
¹⁵⁵ Això passa sovint a l'àmbit públic. Moltes lleis que ordenen determinats sectors d'activitat social estableixen mecanismes d'inspecció i un règim sancionats. De vegades, però, aquests mecanismes simplement no s'activen per manca de recursos (contractar inspectors) o de voluntat política (impopularitat). Hem de reflectir al sistema de classificació aquestes activitats d'inspecció que, encara que contemplades per la llei, no es duen a terme ni se sap quan es farà? Aquest és un problema que l'analista haurà de gestionar amb tacte i respecte a la voluntat de l'organització.

Figura 3-9: Compactació del model funcional



Un cop eliminats els objectius i els projectes, es procedirà a relacionar cada entrada amb una entrada de la columna anterior.

Figura 3-10: Establiment de relacions entre processos



En finalitzar aquesta fase de la construcció del mapa s'ha obtingut una estructura monojeràrquica, amb un sol node pare per cada node fill. És a partir d'aquesta primera estructura jeràrquica, la primera versió del model funcional, que es començaran a buscar les relacions entre els processos, d'acord amb una casuística variada:¹⁵⁶

- Podem trobar nodes que tinguin més d'un node pare. Un exemple seria si hem recollit com a activitat la tramitació de subvencions, i com a funcions la promoció del teatre i la promoció de la música. Ambdues promocions es duen a terme concedint subvencions, per la qual cosa l'activitat de subvencionar depèn dels dos nodes pare. Sobre el model funcional s'establiran en aquest cas dues relacions de tipus estructural o de comprensió.
- Podem trobar també la situació contrària: una activitat que no estigui relacionada amb cap funció. Per exemple, podem localitzar mitjançant l'anàlisi d'un formulari l'activitat de donar d'alta associats a una associació, però no tenir cap menció en la llista de funcions a la gestió d'associats. En aquest cas, la solució adoptada serà crear la funció o macroprocés corresponent a aquesta activitat o activitats, utilitzant eufemismes del tipus "Gestió de".

¹⁵⁶ Per als diferents tipus de relacions que s'inclouran en el mapa, i la forma de descriure-les, vegeu l'apartat 3.2.4.4 d'aquest capítol.

- També podem trobar funcions que no tinguin activitats ni transaccions depenents. Això pot ser símptoma d'una duplicitat no detectada, o bé d'entrades equivalents a objectius, que l'organització no ha desenvolupat funcionalment. En aquest cas, aquestes funcions es marcaran per valorar-ne la supressió durant les revisions iteratives.
- I finalment, podem detectar relacions múltiples contradictòries. Per exemple, imaginem que sabem amb certesa que hi ha una única línia de subvencions a companyies de teatre a nivell d'activitat, la qual podem adscriure a dues funcions documentades: promoció del teatre, i dinamització de la vida teatral del municipi. En aquest cas no podem replicar l'entrada de subvencions sota les dues funcions, sinó que hem de repensar la vigència d'ambdues funcions al sistema de classificació, no ja com a duplicitats sinó com a una estratègia general d'estructuració del sistema de classificació. Això passa sovint quan es concedeix una excessiva importància als plans estratègics com a eix articulador del mapa de processos: el resultat és ràpidament acceptat pel personal de l'organització, però després resulta de confusa aplicació i dificulta l'escalabilitat posterior. És convenient doncs evitar les funcions de denominació retòrica, i assegurar que totes les funcions siguin concretes i fonamentades en activitats clares i de límits identificables.

3.2.4.2.4 Validació mitjançant iteracions

El model funcional és l'instrument ideal per iniciar el procés iteratiu mitjançant el qual totes les persones de l'organització puguin validar i acceptar la forma final del mapa de processos. Aquest mapa, en la forma final, ha de ser familiar i fàcilment comprensible per a tots els membres de l'organització, que no només han de compartir la seva lògica estructural, sinó que també han de sentir-se còmodes amb la seva utilització. Atès que no té encara una forma fixada ni una codificació definitiva, el model funcional no condiciona la interpretació del mapa de processos per part dels membres de l'organització, no permet l'apropiació de funcions per determinades àrees de l'organització (evitant biaixos cap a un sistema orgànic), ni genera rebuig en els usuaris en no tractar-se d'un producte tancat.

Aquest és el moment en què es pot mostrar als membres de l'organització el mapa de processos per a la seva discussió i aprovació. Mostrar-ho abans pot desorientar-los per la seva incompleció, i mostrar-ho més tard pot generar rebuig per la seva aparença de producte definitiu. És també el moment en què l'analista disposa del màxim d'informació sobre les funcions i processos, i d'una estructura madura que permet justificar cada proposta de procés i fonamentar-la mitjançant la referència a les fonts.

A partir d'aquí l'analista es tornarà a reunir amb les àrees i nivells amb què ja es va reunir durant la recopilació d'informació per mostrar-los el mapa, exposar-los els seus dubtes i punts sobre els quals prendre una decisió, recollir els seus suggeriments i aportacions, i obtenir l'acceptació definitiva . En la majoria de casos caldrà més de dues reunions per arribar a aconseguir aquesta acceptació.

3.2.4.2.5 Codificació i denominació definitiva

L'última fase del procés d'elaboració del mapa de processos consisteix a donar al mapa la forma final com a instrument de treball per a tots els membres de l'organització.

El primer element que cal ajustar és la denominació de les entrades. Durant el procés d'elaboració del mapa és preferible mantenir-les en infinitiu i en la forma més desenvolupada per evitar confusions. Però per a la presentació final, és recomanable substantivar aquestes entrades i reduir-ne la longitud a un títol significatiu però prou breu, especialment tenint en compte que la seva interpretació s'enriquirà amb la presència explícita del context en el marc de relacions del mapa.

Així, per exemple, processos com “Tramitar subvencions a companyies de teatre novells” es podria convertir en “Tramitació de subvencions a companyies novells” si depèn d'un macroprocés com “Promoció del teatre”.

Tot i això, es recomana mantenir sempre el verb, preferentment en la seva forma substantivada, ja que la seva eliminació podria generar importants confusions. Així doncs, és millor mantenir la forma “Tramitació de subvencions a companyies novells” que no pas reduir-ho a “Subvencions a companyies novells”. Encara que el seu significat sembli obvi, amb aquest títol no podem determinar si l'activitat consisteix a tramitar subvencions, planificar subvencions o avaluar-ne els resultats.

3.2.4.3 Com es caracteritzen els processos

Els processos són entitats complexes que, dins el mapa global de l'organització, esdevenen el punt d'ancoratge per a la resta d'entitats (sèries, sistemes, etc.), atès que són el que els dona context i sentit dins l'organització. Per aquest motiu la seva descripció es realitza a diferents nivells.

En el moment en què s'identifiquen els processos i es construeix el mapa de processos, s'extreuen els atributs que actuen a nivell de procés. És el que hem anomenat **categorització intrínseca**. Un cop completada aquesta etapa, es pot aprofundir en els nivells inferiors, i identificar i categoritzar aquelles entitats que componen el procés o en són el resultat. Aquesta és la que anomenem **categorització per descomposició**. En

aquesta guia es descriu el primer nivell, la categorització intrínseca, mentre que en la **GDAD** es descriu la categorització per descomposició.

3.2.4.3.1 Caracterització intrínseca

El primer nivell de caracterització del procés és a través de la seva naturalesa funcional intrínseca. Els atributs per a descriure el procés són els que s'indiquen en el model de dades per a l'entitat **ENT002 Procés**:

- **ENT002-001 Identificador**
- **ENT002-002 Nom**
- **ENT002-003 Descripció**
- **ENT002-004 Tipus**
- **ENT002-005 Família**
- **ENT002-006 Àmbit**
- **ENT002-007 Tipus iniciació** (específic CEB)¹⁵⁷
- **ENT002-008 Tipus finalització** (específic CEB)
- **ENT002-009 Tipus recurs** (específic CEB)
- **ENT002-010 Requisits a acreditar** (específic CEB)
- **ENT002-011 Terminis rellevants** (específic CEB)
- **ENT002-012 Necessitat de pagament** (específic CEB)
- **ENT002-013 Data creació**
- **ENT002-014 Data supressió**

Detallem a continuació els dos atributs categòrics o més difícils de determinar.

El primer és el tipus de procés (atribut **ENT002-004 Tipus**). S'inclou com atribut de caracterització la tipologia de processos més utilitzada habitualment, i ja explicada a l'estat de la qüestió, entre processos estratègics, clau i de suport. Però més enllà d'una atribució intuïtiva, aquesta atribució s'intentarà objectivar en funció de qui sigui el destinatari del resultat del procés. Així doncs:

¹⁵⁷ Aquells atributs on s'indica "específic CEB" s'han inclòs en el model com a resultat de necessitats específiques de gestió de la informació del Consorci d'Educació de Barcelona, tal com s'indica en el capítol 4, i el seu ús es considera opcional en qualsevol aplicació d'aquesta guia. En el cas concret dels processos, aquests atributs poden ser rellevants en cas de mapar els processos d'una administració o organisme públic.

- a) Si un procés té com a destinatari directe un usuari o client de l'organització, serà considerat un procés clau.
- b) Si un procés té com a destinatari directe una unitat o equip de treball propi de l'organització, i determina el comportament procedimental o el resultat que tindrà aquesta unitat o equip de treball, serà considerat un procés estratègic.
- c) Si un procés té com a destinatari directe una unitat o equip de treball propi de l'organització, i NO determina, sinó que facilita, el comportament procedimental d'aquesta unitat o equip de treball, serà considerat un procés de suport.

l el segon és la família del procés (atribut **ENT002-005 Família**). Com s'ha comentat abans, es tracta d'una atribució laxa realitzada únicament amb la finalitat de fer agrupacions temàtiques o per àmbits de gestió dels processos. És una herència del costum tradicional d'estructurar els mapes de processos a partir de relacions d'inclusió, creant una jerarquia o taxonomia. S'atribuirà doncs aquest valor de família com a recurs per facilitar la gestió del mapa, i es definiran les famílies a partir dels àmbits principals de gestió de l'organització, o dels objectes finals dels processos (classificació temàtica per clients de procés).

3.2.4.3.2 Caracterització per descomposició

La caracterització d'un procés implica també la descomposició del procés en tots els elements que el componen. Aquests elements, a part dels agents que hi intervenen, són les tasques que realitza cada agent dins del procés, les evidències que rep o genera l'execució de cada tasca, i els elements que es relacionen amb cada evidència, com són els models de document per a les evidències tipificades, les dades que componen cada evidència, els sistemes d'informació que les contenen i les visualitzacions de dades. Aquesta descomposició del procés en els seus elements es detalla en la **GDAD**.

3.2.4.4 Com es relacionen els processos entre ells

Els processos es relacionen entre ells a través de la relació **REL008 procés amb procés**. Aquesta és una relació amb atributs, atès que els processos poden mantenir entre ells els següents tipus de relació:

- **Relació de continuïtat.** L'activitat que finalitza un procés pot, opcionalment, causar (donar lloc a) una activitat d'inici d'un altre procés. D'aquesta manera, el procés que finalitza s'anomena procés precedent, i el procés que comença s'anomena procés conseqüent. El caràcter d'precedent o conseqüent d'un

procés ve determinat, doncs, pel lloc que ocupa en aquesta relació. Això es recull en el valor de l'atribut **REL008-001 Relació de continuïtat**:

- **Procés antecedent**: indica que el procés origen té lloc abans que el procés destí, o com a causa d'aquest, i que la darrera acció o tasca del procés origen determina la primera acció o tasca del procés destí.
 - **Procés conseqüent**: indica que el procés origen té lloc després que el procés destí, o com a conseqüència d'aquest, i que la primera acció o tasca del procés origen ve determinada per la darrera acció o tasca del procés destí.
- **Relació interseccional** (o transversal). Una activitat intermèdia d'un procés pot formar part al mateix temps d'un altre procés. Aquesta relació també la recollirem amb l'atribut **REL008-001 Relació de continuïtat**, on ho indicarem com a **Procés dependent** (indica que el procés origen té lloc de forma simultània o concurrent amb el procés destí, i que una acció o tasca intermèdia del procés origen determina, o ve determinada per, una acció o tasca del procés destí). Segons la forma com es connectin ambdós processos, parlarem de:
 - **Procés transversal**. S'esdevé quan una activitat d'un procés provoca una seqüència d'activitats que segueix un curs diferent, o bé coincideix amb una activitat que ja s'està desenvolupant en un procés diferent. Aquesta és pròpiament la relació transversal entre processos.
 - **Procés derivat o provocat** (anada i tornada), també dit subprocés. S'esdevé quan una activitat d'un procés provoca l'inici d'un altre procés, que acaba retornant al procés que l'ha disparat. L'activitat d'inici i de fi són, doncs, compartides per ambdós processos. En aquest cas podríem dir que es tracta d'un "fals procés", atès que obeeix sovint a convencions de representació (simplificar la representació), convencions de competència (no representar explícitament l'activitat d'agents secundaris o diferents), o a la pertinença del subprocés a un macroprocés diferent al del procés inicial.
 - **Relació de comprensió**. Com hem detallat en l'apartat 3.2.4.2, el concepte de procés o activitat pot descompondre's, a la seva vegada, en altres activitats més petites, exactament amb les mateixes característiques i comportament, fins arribar gairebé a activitats atòmiques. Aquesta flexibilitat fa que es pugui treballar amb diferents nivells de concreció en la identificació dels processos, i que sovint el processos s'agrupin en funcions, macroprocessos o famílies de processos. Aquesta és doncs la relació que es deriva directament del model funcional, una relació habitualment monojeràrquica, que recollirem, un cop estabilitzat el model funcional, en l'atribut **REL008-002 Relació de comprensió**.

3.2.4.5 Com es relacionen els processos amb altres entitats

La primera relació del procés és amb les **disposicions normatives** que regulen qualsevol aspecte del procés, sigui aquest els agents que hi intervenen (regulació competencial), la forma d'esdevenir del procés (regulació procedimental), les eines que ha d'utilitzar el procés (regulació instrumental) o aspectes del resultat del procés (regulació qualitativa).¹⁵⁸ Aquest vincle es recull en la relació **REL010 procés amb disposició normativa**, i dona origen a una entitat complementària, l'entitat **ENT010 Disposició normativa (o mandat)**, que es caracteritza amb els següents atributs:

- **ENT010-001 Identificador**
- **ENT010-002 Títol**
- **ENT010-003 Rang**
- **ENT010-004 Àmbit**
- **ENT010-005 Data vigència**
- **ENT010-006 Data derogació**

La segona relació del procés és amb els **agents**, que en relació amb el procés poden adoptar diferents rols. Aquests rols es materialitzen en:

1. El **responsable**, també dit promotor, del procés. Habitualment correspon a l'agent que té una participació més activa i freqüent en el procés, el qual habitualment és també responsable directe de la qualitat o característiques del resultat del procés. El responsable directe del procés queda recollit en l'atribut **REL002-001 Rol de l'agent** de la relació **REL002 agent amb procés**. En el seu nivell superior (l'Administració o ens al qual pertany aquest responsable), aquesta vinculació queda recollida en l'atribut **ENT002-006 Àmbit del procés** de l'entitat **ENT002 Procés**.
2. El **destinatari** del producte, servei o resultat del procés. Aquest destinatari sempre serà un agent, prèviament identificat en el mapa d'agents, o bé reconegut durant la construcció del mapa de processos, i pot ser intern o extern a l'organització. El destinatari del procés queda recollit també en l'atribut **REL002-001 Rol de l'agent** de la relació **REL002 agent amb procés**.

¹⁵⁸ El concepte de disposició normativa és especialment rellevant en el sector públic, on el principi de competència requereix la constància escrita de qualsevol atribució d'activitat en una norma amb un determinat rang. Això no obsta per a que en l'àmbit privat, malgrat no sigui en tots els casos exigible, siguin també rellevants les normes reguladores (internes o externes) de l'execució dels processos.

3.2.4.6 Com es gestionen les altes, baixes i modificacions de processos

Els processos són una entitat difusa, que és producte d'una identificació feta en un moment i context determinats. En la identificació dels processos sovint pesen factor relacionats amb l'estructura orgànica o competencial, amb objectius estratègics o polítiques puntuals, o amb estils de direcció. Això fa que la variabilitat en els processos sigui difícil de sistematitzar en un conjunt de regles. La pauta que és proposa és per tant purament aproximativa.

Taula 3-2: Pautes d'actuació per al manteniment del mapa de processos

ESDEVENIMENT	ACTUACIÓ
Canvi en les competències assignades al promotor o responsable del procés	Si el procés segueix vigent i el canvi afecta només el promotor o responsable, s'actualitzarà la relació REL002 agent amb procés per reflectir el nou vincle.
Canvi en les famílies de processos	El vincle d'un procés amb una família no és determinant. Un canvi en l'estructura de famílies de processos només implicarà una modificació de l'atribut ENT002-005 Família .
Canvi en l'objecte o finalitat del procés	Si aquest canvi és una mera ampliació (per exemple nous destinataris), només s'actualitzarà la relació REL002 agent amb procés . Però si el canvi implica una modificació total de l'objecte o finalitat, es donarà per caducat el procés informant la data a l'atribut ENT002-014 Data supressió , i es donarà d'alta un nou procés.

ESDEVENIMENT	ACTUACIÓ
Canvi en el procediment o flux d'execució del procés, incloent l'inici i/o la finalització	Els canvis procedimentals són la principal font d'alta i baixa de processos. Sovint provenen de processos de reenginyeria, o de relectures de la identificació dels processos causades per la creació de nous processos i la detecció de noves relacions de dependència (intersecció), de continuïtat o de comprensió entre processos. En aquests casos, la modificació pot comportar la segregació d'un conjunt de tasques d'un procés per a crear un nou procés, la partició d'un procés en dos de nous, la fusió de processos existents, o la baixa d'un procés i la substitució per un de nou. En aquest cas, les modificacions poden afectar gairebé tots els atributs i relacions de l'entitat ENT002 Procés .
Canvis en els sistemes d'informació que donen suport al procés	Els sistemes d'informació es modifiquen amb freqüència. Un canvi en els sistemes d'informació no afectarà la naturalesa del procediment, que continuarà vigent. El vincle amb els sistemes s'actualitzarà creant noves evidències amb l'entitat ENT004 Evidència .
Canvi en la normativa reguladora del procés	Si el canvi normatiu no afecta l'objecte o finalitat del procés, ni tampoc el seu procediment de forma substancial, es limitarà a la creació de nous vincles a través de la relació REL010 procés amb disposició normativa .

3.2.5 Pas 3: mapatge dels sistemes d'informació

Figura 3-11: Estructura per al mapatge dels sistemes d'informació



3.2.5.1 Què s'entén per sistema d'informació

El concepte de sistema d'informació pot contemplar des de la visió més àmplia i neutra (un conjunt d'elements, que pot incloure persones, processos, instruments, etc., que gestiona informació en una organització) fins a una infraestructura o eina tecnològica concreta. Aquesta segona accepció és la que utilitzarem en el mapa de sistemes.

Però no qualsevol element tecnològic serà considerat un sistema d'informació. Les arquitectures tecnològiques de les organitzacions són complexes, i els seus components s'utilitzen per a múltiples finalitats (aplicacions) i són compartits o reutilitzats en múltiples implementacions. Per aquest motiu, i als efectes de mapatge, acotem el concepte de sistema d'informació als següents límits:

- Es tracta d'una eina informàtica, que es pot equiparar a una aplicació o a una base de dades.
- Que emmagatzema informació en format estructurat (dades), tot i que també pugui contenir-ne en format no estructurat (documents-objecte).
- Que té uns límits convencionalment identificables vinculats a una finalitat o a una significació funcional concreta.

3.2.5.2 Com s'identifica un sistema d'informació

La identificació d'un sistema d'informació planteja habitualment molts dubtes. Proposem a continuació un doble criteri per facilitar aquesta identificació: fixar l'abast pel tipus de dades que conté, i determinar els límits per les relacions entre els components.

3.2.5.2.1 Determinació de l'abast: quins sistemes s'inclouen

L'objectiu d'elaborar un mapa de sistemes no és tecnològic, sinó informacional. No es pretén conèixer tots els elements tecnològics de l'organització,¹⁵⁹ sinó només aquells elements que contenen informació, juntament amb els flux que traslladen aquesta informació d'un sistema a un altre. Per tant, limitarem l'abast del mapa a aquelles estructures tecnològiques que contenen informació estructurada, i a les connexions que permeten l'intercanvi d'informació entre elles.

Això inclourà doncs les aplicacions de negoci que generen i utilitzen dades estructurades, com també les bases de dades o conjunts de dades quan aquests siguin accessibles de forma independent a una aplicació o entorn de negoci. Quedaran doncs fora del mapa tots aquells agents tecnològics que no emmagatzemen informació (RPA, connectors, etc.) com també aquells entorns que només emmagatzemen documents-objecte (xarxa local, carpetes compartides, etc.).

El motiu d'aquesta limitació de l'abast prové del següent fet. Tots els sistemes d'informació emmagatzemen informació d'alguna manera, siguin documents digitals, dades de negoci, metadades o traces d'activitat. De la mateixa manera, tots els documents que formen part de qualsevol agrupació documental, quan estan en format digital, poden ser emmagatzemats en un o diferents sistemes d'informació (actuant de repositoris), o fins i tot enlloc. Quan la informació és emmagatzemada en forma d'objectes o fitxers (informació discreta), com s'esdevé quan s'utilitza un EDMS, un *file system* o una xarxa local, aquests documents poden ser gestionats com unitats independents, preservant el seu context original, i si cal poden ser desplaçats a un altre sistema sense perdre cap element essencial del seu context arxivístic. Són, per tant, independents del sistema que els hostatja. Però quan aquests documents són emmagatzemats únicament com a dades dins un sistema d'informació estructurada (informació contínua), la seva existència lògica com entitats documentals passa a dependre del sistema que els conté, amb la qual cosa esdevenen captius d'aquest sistema. Extreure aquests documents del sistema per donar-los existència física depèn per tant de les funcionalitats del propi sistema, i requereix habitualment algun tipus de transformació sobre les dades que pot impactar sobre el seu valor arxivístic (com, per exemple, quan es genera un *dataset* a efectes d'arxiu a partir d'un sistema de negoci). Per poder doncs integrar aquests conjunts de dades en el mapa informacional d'una organització es fa necessari, per tant, incorporar en el mapa global els sistemes d'informació que contenen informació estructurada.

¹⁵⁹ Aquesta funció es materialitza habitualment amb un inventari o catàleg de sistemes d'informació i aplicacions del qual disposen la majoria de departaments de tecnologies de la informació a les organitzacions.

3.2.5.2.2 **Determinació dels límits: components del sistema**

Els sistemes d'informació, segons la definició que s'ha exposat anteriorment, poden correspondre a estructures tecnològiques molt diverses. Cada cop més els components tecnològics es reutilitzen per a diferents finalitats, i el que fa una dècada era una aplicació informàtica única ara és un conjunt d'interaccions entre diferents components, que poden residir en ubicacions diferents i sota responsabilitat d'agents diferents. Els escenaris més habituals en els sistemes de negoci d'una organització són:

- **Aplicacions informàtiques que segueixen un model tradicional**, amb uns límits definits i un conjunt de components a nivell d'interfície, processos i dades d'ús exclusiu d'aquella aplicació (per exemple un sistema de gestió de recursos humans basat en SAP).¹⁶⁰ En aquest conjunt s'inclourien també les solucions basades en eines d'oficina (ofimàtiques).
- **Solucions tecnològiques puntuals**, que es basen en la reutilització de components (funcions, macros, serveis Web, etc.), tant a nivell d'interfície com de procés, i que accedeixen a diferents fonts de dades habitualment desvinculades de l'aplicació. Una solució de visualització sobre mapa, que utilitza dades d'un fitxer CSV, cartografia pública en JSON i APIs de *Google* per visualitzar les dades sobre *Google Maps* en seria un exemple.
- **Solucions basades en plataformes de col·laboració**, habitualment en modalitat SaaS, que combinen components diferents de la plataforma per a construir un conjunt de funcionalitats, i que inclouen programació *low code*. Un *site* de Sharepoint que combini llistes, biblioteques, formularis i flux programats seria un exemple d'aquest tipus.

En el primer cas, la identificació és relativament senzilla, atès que el sistema té uns límits clarament reconeixibles, un nom propi i uns components tecnològics exclusius, i és reconegut com a tal per l'organització. Però els altres dos casos són més complexos d'identificar.

El criteri que adoptarem per poder identificar un sistema com a tal en el mapa de sistemes serà el següent:

- A. El sistema dona suport de forma clara a la gestió d'un procés prèviament identificat.
- B. El sistema emmagatzema dades del procés al qual dona suport amb voluntat de permanència temporal.

¹⁶⁰ Programari de gestió de gran difusió en el món empresarial. Per a més informació vegeu <https://www.sap.com/spain/index.html>.

- C. El sistema, per donar suport a la gestió del procés, mobilitza una sèrie de components tecnològics que mantenen entre ells un vincle funcional (es comuniquen i intercanvien informació).

Quan dins una mateixa aplicació o plataforma es puguin identificar diferents sistemes, s'aplicarà com a criteri d'identificació per a la segmentació els diferents processos pels quals s'utilitza el sistema.

Un cop identificat, el sistema d'informació generarà una instància de l'entitat **ENT011 Sistema d'informació o aplicació** del model de dades.

3.2.5.3 Com es caracteritzen els sistemes

La caracterització dels sistemes és una tasca complexa, i requereix disposar d'informació contextual suficient de caràcter tant tecnològic com funcional.

Els atributs identificatius són els següents:

- **ENT011-001 Identificador**
- **ENT011-002 Nom**
- **ENT011-003 Nom curt**
- **ENT011-019 Data inici servei**
- **ENT011-020 Data fi servei**

Les pautes per informar les dates d'inici i de fi de servei es troben descrites en el model de dades, i també a l'apartat "Com es gestionen les altes, baixes i modificacions de sistemes d'informació".

Els atributs de caràcter funcional són els següents:

- **ENT011-004 Responsable funcional**
- **ENT011-005 Gestor TIC** (específic CEB)
- **ENT011-006 Proveïdor**
- **ENT011-010 Abast**
- **ENT011-016 Descripció**
- **ENT011-017 Funcionalitat**

Les pautes per informar aquests atributs es troben detallades en la descripció dels propis atributs en el model de dades (apartat 4.6.1.11 del capítol 4). En el cas de l'abast, aquest vindrà determinat no només per l'àmbit d'actuació, sinó també per la tecnologia

utilitzada, descrita en els atributs **ENT011-007 Entorn** i **ENT011-008 Entorn tecnològic**. Així doncs, les solucions ofimàtiques no es podran considerar corporatives (és a dir pròpies de la voluntat de l'organització) sinó simplement suports a l'usuari final. Les aplicacions, en canvi, veuran determinat el seu abast per la pertinença del seu promotor a un determinat nivell de l'organització.

Finalment, els atributs de caràcter tecnològic són els següents:

- **ENT011-007 Entorn**
- **ENT011-008 Entorn tecnològic**
- **ENT011-009 Model de servei**
- **ENT011-011 Restriccions semàntiques**
- **ENT011-012 Restriccions de format**
- **ENT011-013 Controls d'integritat**
- **ENT011-014 Controls d'accés**
- **ENT011-015 Classificació de seguretat** (específic CEB)
- **ENT011-018 Ubicació** (específic CEB)

Les pautes per informar aquests atributs es troben detallades també en el model de dades (apartat 4.6.1.11 del capítol 4).

A l'hora de determinar l'entorn, un dels atributs clau als efectes de classificació dels sistemes, cal tenir present que qualsevol solució tecnològica pot implicar un nivell d'esforç considerable en el seu desenvolupament, i pot posar-se a l'abast de múltiples usuaris. Per tant, la tecnologia utilitzada i la disponibilitat de la solució o aplicació per al conjunt de l'organització no són factors discriminants per determinar si una solució es pot considerar ofimàtica o aplicació. Per tant, ens guiarem en aquest cas pels nivells d'administració:

- a) Aquelles solucions on l'usuari de l'aplicació sigui al mateix temps l'administrador es consideraran solucions ofimàtiques.
- b) Aquelles solucions on existeixi un administrador que pugui controlar l'accés dels usuaris i la compartició i ús de la informació es consideraran aplicacions.

El fet que una solució sigui ofimàtica o aplicació determinarà també si incorpora controls d'accés i d'integritat de les seves dades. Habitualment tota aplicació inclourà com a mínim un control d'accés, i molt probablement també un control d'integritat, mentre que en les solucions ofimàtiques aquests controls seran molt poc freqüents.

3.2.5.4 Com es relacionen els sistemes entre ells

En el cas dels sistemes d'informació, les relacions que contempen en aquesta metodologia es tradueixen en intercanvis d'informació, que poden ser de dos tipus:

1. Intercanvi d'instruccions
2. Intercanvi de dades

L'intercanvi d'instruccions afecta preferentment a l'anàlisi de processos, mentre que l'intercanvi de dades afecta sobretot a l'anàlisi de sèries i circuits documentals. Als efectes de mapatge de sistemes ens centrarem en les relacions d'intercanvi de dades. Aquestes relacions poden de diferents tipus en funció del moment i de la forma.

Pel que fa al moment, poden ser:

1. En temps real.
2. En temps diferit (programat).
3. En temps diferit (ocasional o invocat).

Pel que fa a la forma, poden ser:

1. Intercanvi sense transformació de la informació.
2. Intercanvi amb transformació automatitzada de la informació (inclou ETL).
3. Intercanvi amb transformació manual (humana) de la informació.

Els intercanvis d'informació entre sistemes, més enllà de les relacions lògiques en la fase de modelat, es tradueixen en connexions vinculades a una implementació concreta. Aquestes connexions no són de caràcter general, sinó únicament de caràcter instrumental, i es desenvolupen com a resultat d'una petició/acció voluntària de l'organització, no de forma natural ni orgànica. La lògica que les justifica és bàsicament funcional, però també d'oportunitat, atès que sovint és materialment impossible (per motius tecnològics o pressupostaris) implementar totes les connexions entre sistemes que serien desitjables per tal d'assolir el principi de la dada única.¹⁶¹

Per recollir en el mapa de sistemes les relacions entre ells, en aquest model s'ha creat una entitat específica que permet descriure amb detall les característiques de cada intercanvi de dades. Es tracta de l'entitat **ENT012 Interfície de connexió entre sistemes**, que correspon a la relació:

¹⁶¹ Aquest és, precisament, l'objectiu d'una política de govern de la dada: aconseguir connectar tots els sistemes d'informació entre ells i sincronitzar-ne totes i cada una de les dades que contenen (Benfeldt et al., 2020).

- **Sistema d'informació** es pot comunicar amb un o més **sistemes d'informació** a través d'una o més **interfícies**

En el context d'aquest model anomenem doncs interfície a qualsevol procés o flux de dades entre dos sistemes d'informació. La interfície pot incorporar algun element tecnològic (per exemple un connector, un codi de procés ETL, etc.), o bé limitar-se a un procés d'exportació-importació de dades. En tota interfície sempre hi ha un sistema que és origen de les dades, un sistema destí, unes dades que viatgen i unes condicions d'aquest viatge. El concepte d'interfície d'aquest model correspon en l'àmbit de la ciència de dades al concepte de **flux de dades** (*dataflow*).¹⁶²

Els atributs que descriuen una interfície són els següents:

- **ENT012-001 Identificador**
- **ENT012-002 Nom**
- **ENT012-003 Tipus d'interfície**
- **ENT012-004 Tipus d'importació**
- **ENT012-005 Procediment d'importació**
- **ENT012-006 Actualització**
- **ENT012-007 Historificació**
- **ENT012-008 Nom del fitxer de càrrega** (específic CEB)
- **ENT012-009 Restriccions semàntiques**
- **ENT012-010 Restriccions de format**
- **ENT012-011 Observacions**
- **ENT012-012 Data inici servei**
- **ENT012-013 Data fi servei**

Dos grups d'atributs mereixen un detall especial. El primer és el que correspon a la forma com es realitza la connexió entre els sistemes, que es descriu a partir de tres atributs:

- L'atribut **ENT012-003 Tipus d'interfície** descriu de quina manera s'obtenen les dades del sistema origen per a comunicar-les al sistema destí. Aquesta obtenció pot ser totalment automatitzada, o pot requerir diferents tipus d'intervenció humana.
- L'atribut **ENT012-004 Tipus d'importació** descriu la forma com es traslladen les dades d'un sistema a un altre. Parteix de l'opció en la qual les dades no es mouen

¹⁶² Pel concepte vegeu Julbe, 2023.

del sistema origen, sinó que el sistema destí hi pot accedir quan n'hagi de menester. La segona opció és que ambdós sistemes poden intercanviar dades mitjançant algun tipus de connexió. I la tercera s'esdevé quan un agent extern accedeix al primer sistema, descarrega les dades, i les carrega posteriorment al sistema destí.

- Finalment, l'atribut **ENT012-005 Procediment d'importació** classifica el procediment d'importació en funció del grau de tractament i operacions que cal realitzar sobre les dades durant la transferència, i també segons el nivell d'intervenció dels agents humans.

I el segon grup d'atributs és el que correspon als moments o periodicitat en què:

- S'executa el flux que trasllada les dades del sistema origen al sistema destí (atribut **ENT012-006 Període d'actualització**).
- S'executa el flux que genera la versió d'arxiu (en el sistema destí) de les dades del sistema origen (atribut **ENT012-007 Període d'historificació**).

3.2.5.5 Com es relacionen els sistemes d'informació amb altres entitats

La primera relació que cal cercar consisteix en identificar el promotor o responsable del sistema. De forma general, es considerarà promotor del sistema aquell agent que és al mateix temps propietari del procés al qual dona suport el sistema. Si el sistema dona suport a processos amb diferent propietari, es considerarà promotor del sistema al nivell orgànic més alt de l'organització. No es considerarà promotor al desenvolupador del sistema, sigui aquest una àrea de tecnologia interna o un proveïdor extern, atès que el promotor del sistema serà també el responsable de les dades que contindrà el sistema. El promotor coincideix habitualment amb l'agent que ha facilitat els requeriments funcionals per al desenvolupament del sistema, i que en serà posteriorment l'usuari principal o administrador funcional. Aquesta relació és la que es recull a la relació **REL006**:

- **Sistema d'informació** és promogut per un **agent**

La relació que s'estableix entre un procés i un sistema d'informació ve determinada per les funcionalitats del sistema, és a dir les activitats del procés que pot recollir, i si escau automatitzar. Aquestes relacions poden ser:

1. **Interactives**. Un usuari humà interacciona amb el sistema a través d'interfícies, sensors, etc.

2. **Autònomes o desateses.** El propi sistema executa accions predefinides a partir de regles o instruccions, en nom d'un determinat agent. Això inclou la presa de decisions basades en intel·ligència artificial.

En aquest model, les relacions del procés amb el sistema d'informació es representen en forma d'interaccions del procés, i en conseqüència de les evidències que materialitzen aquestes interaccions (formularis, pantalles de visualització de dades, generació automatitzada d'evidències, etc.). Per tant quedaran recollides en dues relacions consecutives, la relació **REL016**:

- **Procés** genera una o més **evidències**

I després en la relació **REL019**:

- **Evidència** està continguda en un **sistema d'informació** (quan és de tipus "dades")

3.2.5.6 Com es gestionen les altes, baixes i modificacions de sistemes d'informació

La variabilitat en la vida dels sistemes d'informació va aparellada a la dificultat en la seva identificació. Els sistemes poden evolucionar, afegir components o funcionalitats, patir canvis en l'estructuració de les seves dades, etc. Intentem a continuació recollir de forma sintètica aquesta casuística, i la forma de procedir en cada cas:

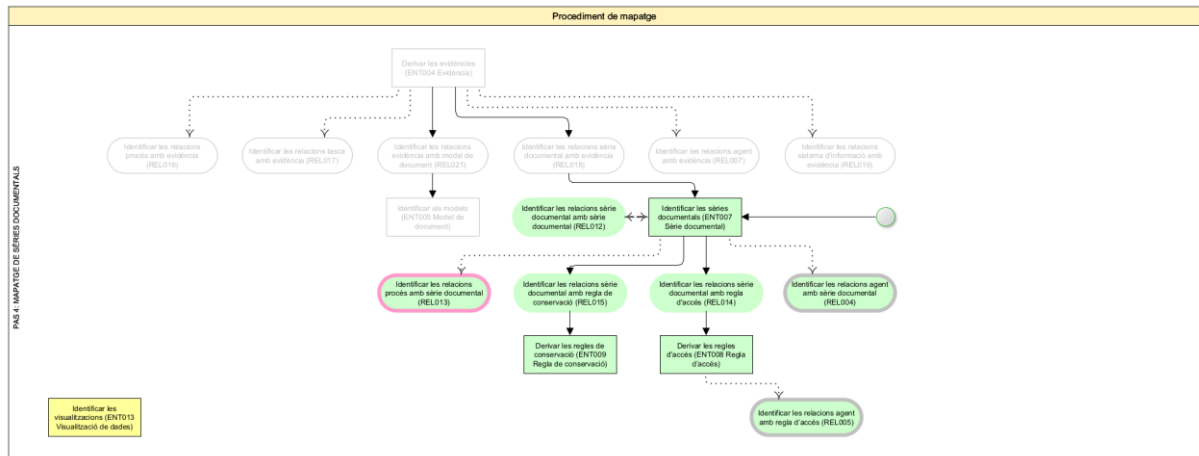
Taula 3-3: Pautes d'actuació per al manteniment del mapa de sistemes d'informació

ESDEVENIMENT	ACTUACIÓ
Fi de la utilització d'un sistema d'informació	<p>Quan un sistema d'informació deixi de funcionar de forma definitiva, i deixi d'estar disponible en un entorn productiu, s'actualitzarà l'atribut ENT011-020 Data de fi de servei.</p> <p>Aquesta finalització del servei només afecta les funcionalitats, no la informació continguda en el sistema, que pot continuar operativa un cop migrada a un nou sistema, o bé establitzada dins el sistema finalitzat en forma de dades llegades (<i>legacy system</i>).</p>

ESDEVENIMENT	ACTUACIÓ
Canvi de versió en la tecnologia d'un sistema d'informació	Els canvis de versió que no impliquin canvi en la tecnologia de base o en els components del sistema no seran reflectits en el mapa.
Canvi de tecnologia d'un sistema d'informació	<p>Quan un sistema d'informació sigui migrat a un nou entorn tecnològic, o incorpori components tecnològics no existents anteriorment, es farà constar el nou entorn a l'atribut ENT011-008 Entorn tecnològic, i si escau es descriurà el canvi a l'atribut ENT011-017 Funcionalitats. Si és necessari també s'actualitzarà l'atribut ENT011-018 Ubicació, com també qualsevol altre atribut que es vegi afectat pel canvi tecnològic (ENT011-011 Restriccions semàntiques, ENT011-012 Restriccions de format, ENT011-013 Controls d'integritat, ENT011-014 Controls d'accés, etc.).</p>
Fusió de dos sistemes d'informació	En el cas de fusió de dos sistemes d'informació, es farà constar la finalització d'activitat d'ambdós sistemes, i es donarà d'alta un nou sistema.

3.2.6 Pas 4: mapatge de les sèries documentals

Figura 3-12: Estructura per al mapatge de les sèries documentals



3.2.6.1 Què s'entén per sèrie documental

Si el mapa de processos permet modelar una vista funcional de l'organització, el mapa de sèries documentals és l'eina que permet operativitzar aquest modelat sobre un conjunt d'evidències, sistemes d'informació i dades concrets. La sèrie documental s'ha d'entendre doncs com una entitat física, creada a efectes pràctics per agrupar informació a partir d'un modelat funcional previ i poder-hi aplicar conjunts d'accions de tractament mitjançant regles que s'activen al llarg del seu cicle de vida.

Tradicionalment la sèrie documental s'ha definit com l'agrupació de tots els expedients que comparteixen un mateix tipus de tramitació o un mateix tipus de tema o subjecte al que fan referència, i s'ha incardinat en diferents nivells dels nombrosos models de desplegament jeràrquic de les agrupacions documentals.¹⁶³ Però en aquesta metodologia adoptem un plantejament diferent.

Com s'ha detallat en l'apartat 2.3.1, i especialment en el 2.3.1.3, del capítol 2,¹⁶⁴ el concepte de sèrie documental té dues dimensions: la dimensió funcional i la dimensió orgànica.

La **dimensió funcional** prové de la relació que el concepte de sèrie manté amb una determinada funció, activitat o procés de l'organització on identifiquem aquesta sèrie.

¹⁶³ Vegeu el model de dades per a les aportacions de les principals normes de descripció arxivística en aquest sentit.

¹⁶⁴ Apartats dels quals a continuació en repetim literalment algunes idees per facilitar la comprensió d'aquesta part de la guia.

Segons aquesta dimensió, la sèrie documental és tota aquella documentació o conjunts d'evidències, en qualsevol forma i suport, que genera una organització en el desenvolupament d'una funció o activitat pròpia.

La **dimensió orgànica** prové de la relació que la sèrie manté amb l'òrgan que la genera. Aquesta dimensió intersecciona amb la funcional, de forma que ens permet definir la sèrie documental com aquella documentació, en qualsevol forma i suport, que genera un determinat òrgan d'una organització en el desenvolupament d'una funció o activitat pròpia.

Situant la dimensió funcional en l'eix d'abscisses, i la dimensió orgànica en l'eix d'ordenades, obtenim una matriu que ens permet identificar totes les sèries d'un organisme, i al mateix temps posar-les en relació tant per la dimensió funcional com per l'orgànica. Aquesta matriu és el que, en el marc d'aquesta metodologia, s'anomena **mapa de sèries documentals**.

Així doncs, per poder construir el mapa de sèries documentals, tota organització necessita disposar prèviament de dos elements:

- Una relació de tots els agents de l'organització que poden actuar com a productors de sèries documentals. Aquest instrument és el que en el marc d'aquesta metodologia anomenem **mapa d'agents** (vegeu apartat 3.2.3).
- Una relació de totes les funcions, activitats i processos que pot portar a terme l'organització en l'exercici de les seves competències o interessos. Aquesta relació en el marc d'aquesta metodologia l'anomenem **mapa de processos** (vegeu apartat 3.2.4).

El principal potencial del mapa de sèries documentals és que facilita l'anàlisi transversal del substrat documental de l'organització, ja que permet posar en relació documents formalitzats, sistemes d'informació, i evidències d'altres tipus, abstraient la forma i l'estructura concretes de cada document de la seva vinculació amb una activitat i un cicle de vida específics. D'aquesta manera, la sèrie documental es converteix en l'eix per a la integració de la gestió documental d'una organització.

El concepte de sèrie documental és també un element indispensable per a l'anàlisi i el redisseny de processos. Quan s'analitza el circuit documental d'una tramitació, abans de plantejar propostes de millora cal conèixer amb detall el substrat documental vigent, la composició, l'abast i les deficiències. En la majoria dels casos, són diverses les sèries documentals que es veuen afectades per cada procés. El mapa de sèries documentals permet analitzar i simular l'impacte que el redisseny d'un procés pot suposar sobre aquestes sèries documentals i els sistemes d'informació que els donen suport, i calcular el cost i els beneficis d'implementar determinades accions de millora. En la **GDAD** es desenvolupa aquest vincle entre les sèries documentals i la millora de la qualitat dels processos (apartats 3.3.2 i 3.3.3 d'aquest capítol).

3.2.6.2 Com s'identifica una sèrie documental

La sèrie documental es pot identificar durant l'aproximació que es descriu en aquesta guia, i també a partir de l'aproximació que es descriu en la **GDAD**. Tant en un cas com en l'altre el raonament a seguir és el que s'indica a continuació:

3.2.6.2.1.1 Identificació de la sèrie primària teòrica

Per a la identificació de les sèries d'una organització s'aplicarà la següent regla: per cada procés identificat en el mapa de processos, es crearà una sèrie documental que tindrà com a productor al promotor o responsable d'aquell procés. Aquesta és la sèrie que sempre existirà, i que tindrà la consideració de **sèrie primària**.

La sèrie primària és la sèrie de referència de tot procés, i aquella a partir de la qual es podran derivar altres sèries dins el mateix procés, o es podran establir relacions amb sèries d'altres processos. Per un detall de les relacions entre les sèries documentals, vegeu l'apartat 3.2.6.4 d'aquest capítol.

3.2.6.2.1.2 Identificació de les sèries reals

La sèrie documental com a intersecció ofereix sempre una única sèrie per procés i agent, el que en denominem sèrie primària. Però habitualment la realitat és més complexa, i la forma com s'han documentat els processos no correspon amb aquest model teòric. Per poder emprendre accions de millora, és necessari que el mapa de sèries reflecteixi la realitat documental de l'organització, cosa la qual portarà a la identificació de més d'una sèrie per cada intersecció procés-agent.

La identificació de les sèries reals, més enllà de la sèrie primària teòrica, es basa en el treball de camp, i se centra en el concepte d'expedient. Per identificar totes les sèries generades per un procés es demanarà a tots els agents que participen en el procés que facilitin documents i expedients de mostra, i també que facilitin accessos als sistemes d'informació i aplicacions als quals accedeixen per executar el procés. Amb la informació recollida es crearan agrupacions d'evidències (documents i accessos a sistemes d'informació) corresponents a expedients per agent, i s'analitzaran les seves característiques i vincles.

Això ha de permetre identificar:

- Si existeix una sèrie primària real reconeguda com a tal, o si apareixen més d'una sèrie candidates a primàries. Això s'esdevé en aquells processos on el rol de promotor o responsable no ha estat clarament establert, o bé en aquells processos que s'han definit a un nivell de concreció massa alt. En aquest darrer

cas, l'anàlisi de les sèries pot evidenciar que corresponen a subprocessos clarament separables, amb la qual cosa s'obrirà la possibilitat de subdividir el procés per racionalitzar el seu substrat informacional.

- Si apareixen sèries secundàries, o candidates a secundàries, derivades de la sèrie primària, estiguin o no formalment reconegudes com a tals.
- Si s'utilitzen sistemes d'informació o evidències compartides amb altres processos, cosa la qual permetrà identificar sèries relacionades de diferents tipologies.

En aquest punt l'enfocament *top-down* s'atura, i es limita a reflectir la realitat identificada (un procés i un conjunt de sèries relacionades). Si es volen explotar les possibilitats de millora del procés a través de la millora del seu substrat informacional, caldrà utilitzar l'enfocament *bottom-up* tal com es descriu a la **GDAD**.

3.2.6.3 Com es caracteritzen les sèries documentals

Totes les evidències que componen una sèrie documental, siguin documents-objecte, sistemes d'informació o altres conjunts de dades, comparteixen:

- El mateix vincle amb un procés
- El mateix agent productor
- El mateix context funcional, jurídic, administratiu i de negoci
- El mateix cicle de vida, disposició i calendaris (adaptat a cada tipus d'evidència)
- El mateix patró de relacions amb altres elements del mapa informacional de l'organització

Això és precisament el que fa especialment potent el concepte de sèrie en el tractament de la informació, atès que actua com a recurs reductor d'una realitat habitualment voluminosa i aparentment molt diversa, sotmesa a una elevada subjectivitat en la seva interpretació si és tractada de forma independent. La sèrie documental fa possible identificar evidències de formes aparentment diverses com pertanyents a la mateixa realitat, vincular-les al seu context, i decidir sobre accions determinants del seu cicle de vida. Per aquest motiu, si assignar una evidència a una sèrie li proporciona un context estable, caracteritzar la sèrie equival a definir i homogeneïtzar les característiques de les evidències que en formen part.

La caracterització de la sèrie documental consisteix doncs en la descripció de les característiques del seu expedient tipus. De la mateixa manera que amb els processos,

aquesta descripció es pot realitzar a partir de les característiques intrínseques a la sèrie documental, o bé per descomposició de la sèrie documental.

3.2.6.3.1 Caracterització intrínseca

El primer nivell de caracterització de la sèrie documental és a través de les característiques del seu expedient tipus, les quals defineixen la forma com aquest s'instanciarà. Els atributs per a descriure la sèrie documental són els que s'indiquen en el model de dades per a l'entitat **ENT007 Sèrie documental**:

- **ENT007-001 Identificador**
- **ENT007-002 Nom**
- **ENT007-003 Criteri de nominalització**
- **ENT007-004 Caràcter**
- **ENT007-005 Descripció**
- **ENT007-006 Vital**
- **ENT007-007 100% Digital** (específic CEB)
- **ENT007-008 Data creació**
- **ENT007-009 Data supressió**

3.2.6.3.1.1 Tipificació de la sèrie

El primer element a determinar és la vinculació de la sèrie amb la sèrie primària teòrica. És a dir, cal indicar si la sèrie és candidata a sèrie primària, o bé si es pot determinar ja d'antuvi el seu caràcter clarament secundari.

En aquesta fase es pot atribuir a una sèrie el caràcter de primària, si es considera que és candidata a ser-ho (per exemple perquè el seu productor és el promotor del procés), malgrat que encara no compleixi els requisits de qualitat de tota sèrie primària. Aquests requisits que ha de complir una sèrie primària correctament formada són els següents:

- a) Documenta totes les tasques essencials del procés. Una tasca és essencial quan eliminar aquesta tasca del procés té com a conseqüència aturar el procés, o evita que el procés arribi al seu final esperat.
- b) Totes aquestes tasques es documenten amb evidències originals, és a dir, evidències emanades directament de les tasques que s'estan executant.

Aquestes evidències es poden substituir per còpies quan aquestes còpies estan autenticades i tenen la mateixa qualitat probatòria que l'evidència original.

- c) L'expedient ha de ser conservat pel productor, o en un entorn on el productor pugui tenir control sobre la cadena de custòdia.
- d) Totes les evidències han d'estar enregistrades en un suport que permeti conservar-les enteres i autèntiques durant el temps que estableixi la regla de conservació.

Aquesta classificació es farà constar a l'element **ENT007-004 Caràcter**.

3.2.6.3.1.2 Assignació del criteri de nominalització

El criteri de nominalització es pot definir com la pauta segons la qual s'identificaran de forma unívoca els expedients d'una sèrie documental. Aquesta identificació no és trivial, ni es redueix a la mera assignació d'un codi identificador. El criteri de nominalització determina la granularitat amb la qual se subdividirà la sèrie documental, la qual vindrà determinada pel que s'estableixi com a objecte del procés, i per tant de l'expedient. I no només això: el criteri de nominalització també determina la seqüència amb què s'ordenaran els expedients dins la sèrie documental. Per tant, per assignar el criteri de nominalització cal tenir en compte algunes recomanacions:¹⁶⁵

1. En els processos i sèries que generen expedients reglats, és important fixar com a objecte de l'expedient el nivell més concret possible, aquell sobre el qual s'executa directament el procés, i defugir agrupacions (per exemple, en un procés de beques es recomana nominalitzar per sol·licitud de beca, i no pel sol·licitant o perceptor de la beca).
2. En els processos i sèries que generen expedients no reglats tècnics s'intentarà també, sempre que sigui possible, ajustar l'objecte de l'expedient a l'objecte real i últim del procés (per exemple, en un procés d'organització d'activitats de formació es recomana nominalitzar per activitat).
3. En els processos i sèries que generen expedients no reglats temàtics es podran agrupar els objectes, o substituir-los per agrupacions geogràfiques, cronològiques, etc., tenint present que habitualment els expedients temàtics corresponen a sèries documentals secundàries i amb un termini de conservació curt.
4. En els entorns digitals la consecutivitat en la nominalització d'expedients (l'habitual número d'expedient consecutiu) és irrellevant als efectes de cerca, gestió, i fins i tot d'ordenació, i cal prioritzar les nominalitzacions significatives

¹⁶⁵ Per als tipus d'expedient vegeu l'apartat 2.3.1.4.1 del capítol 2.

(codi d'objecte) per sobre de les nominalitzacions convencionals (número d'expedient consecutiu).

5. Cal defugir també la inclusió com a part del criteri de nominalització d'informacions que ja constaran com a metadades de l'expedient, com per exemple un codi de productor, les inicials del procés o l'any de l'expedient.
6. En cas que es faci referència a un determinat objecte, si aquest objecte disposa d'algun tipus de codificació estandarditzada (per exemple el número de client en un registre de clients), és important que el criteri de nominalització de totes els sèries que actuen sobre aquest objecte utilitzi la mateixa identificació per a l'objecte, intentant evitar que es creïn identificacions diferents per al mateix objecte en sèries diferents.

Aquesta metodologia recomana utilitzar una determinada notació per representar el nom de l'expedient:

- Sempre s'inicia amb el prefix "N-", per indicar, en cas de tractaments massius de les dades, que es tracta d'un nom o número d'expedient.
- El nominal pot contenir parts fixes i parts variables.
- Les parts fixes s'escriuen sense cap notació especial, mentre les parts variables s'escriuen entre claudàtors ("[]").

Així doncs, el criteri per nominalitzar els expedients d'una convocatòria de beques es podria representar de la forma següent:

N-[número convocatòria]/[número sol·licitud]

Els prefix utilitzats per a definir els criteris de nominalització podrien derivar amb el temps en nous criteris de classificació, o servir per a detectar nous processos, i també en termes d'etiquetatge en un vocabulari controlat.

El criteri de nominalització es recollirà a l'atribut **ENT007-003 Criteri de nominalització**.

3.2.6.3.1.3 Determinació de la vitalitat o de la sèrie

Es considera vital aquella sèrie documental la conservació de la qual és determinant per a la supervivència de l'organització i el manteniment de la seva activitat.

Per a determinar la vitalitat d'una sèrie documental, s'analitzarà la informació que conté (a través de les evidències de la sèrie). La sèrie serà vital si es compleixen les quatre condicions següents:

- a) Si la informació és autèntica.

- b) Si la informació no existeix en cap altre sèrie documental antecedent, relacionada o recapitulativa, o hi existeix de forma incompleta.
- c) Si l'objecte descrit per aquesta informació forma part dels objectes de negoci de l'organització (destinatari, clients, etc.) o bé dels seus recursos (personal, patrimoni, etc.).
- d) Si la seva pèrdua podria provocar conseqüències operatives o jurídiques irreparables.

El caràcter de vital d'una sèrie documental es recull a l'atribut **ENT007-006 Vital**.

3.2.6.3.1.4 Determinació del cicle de vida de la sèrie

La determinació del cicle de vida de la sèrie documental (del seu termini de conservació, transferència, transformació, etc.) forma part de la caracterització de la sèrie, i es materialitza amb la instanciació de l'entitat **ENT009 Regla de conservació**, la qual es caracteritza amb els següents atributs:

- **ENT009-001 Identificador**
- **ENT009-002 Esdeveniment d'inici o disparador**
- **ENT009-003 Termini**
- **ENT009-004 Resolució**
- **ENT009-005 Criteri de mostreig**
- **ENT009-006 Suport de destí**
- **ENT009-007 Format de destí**
- **ENT009-008 Lloc de destí**
- **ENT009-009 Data creació**
- **ENT009-010 Data supressió**

La sèrie documental s'hi vincula amb la relació **REL015 sèrie documental amb regla de conservació**.

No és objecte d'aquesta metodologia descriure el procés d'elaboració d'una regla de conservació.¹⁶⁶

¹⁶⁶ Aquest procés disposa d'una norma ISO específica (*ISO/TR 21946:2018 Information and documentation — Appraisal for managing records*), i estan en curs iniciatives en que van en la línia exposada, com n'és exemple la *Comunidad de práctica VALORA* (<https://www.archiverosdeandalucia.org/valora/>) (Casellas, 2019).

3.2.6.3.1.5 Determinació del règim d'accés de la sèrie

La determinació del règim d'accés a la sèrie documental, i de publicitat del seu contingut, forma part també de la caracterització de la sèrie, i es materialitza amb la instanciació de l'entitat **ENT008 Regla d'accés**, la qual es caracteritza amb els següents atributs:

- **ENT008-001 Identificador**
- **ENT008-002 Esdeveniment d'inici o disparador**
- **ENT008-003 Termini**
- **ENT008-004 Resolució**
- **ENT008-005 Motiu de limitació**
- **ENT008-006 Criteri d'aplicació de la regla**
- **ENT008-007 Criteri de publicitat per transparència**
- **ENT008-008 Data creació**
- **ENT008-009 Data supressió**

La sèrie documental s'hi vincula amb la relació **REL014 sèrie documental amb regla d'accés**.

No és objecte d'aquesta metodologia descriure el procés d'elaboració d'una regla d'accés.¹⁶⁷

3.2.6.3.2 Caracterització per descomposició

La caracterització d'una sèrie documental implica també la identificació de totes les evidències que en formen part. Aquestes evidències poden ser documents-objecte, o bé interaccions amb sistemes d'informació o aplicacions, a través de les quals la sèrie es vincula també amb sistemes d'informació. La posterior extracció de les dades elementals contingudes a cada evidència permet descompondre la sèrie fins al nivell de dada, i construir d'aquesta manera la matriu de variables d'informació o dades en què es condensa l'essència evidencial de la sèrie documental. Aquest és el procés que es descriu a la **GDAD**.

¹⁶⁷ La definició de polítiques d'accés està afectada per marcs normatius multisectorials (protecció de dades, transparència, seguretat informàtica, etc.) i resta fora de l'abast d'aquesta tesi.

3.2.6.4 Com es relacionen les sèries entre elles

Les sèries documentals hereten el seu marc de relacions de les relacions entre els processos. Però afegixen a aquestes relacions algunes que són específiques de les sèries documentals.¹⁶⁸

Una part de les relacions entre les sèries ve determinada per la seva pertinença a un **mateix procés**. Això vol dir que la formació de la sèrie origen i destí coincideixen en el temps, i que determinades evidències intermèdies de la sèrie origen poden existir també en la sèrie destí. Si les dues sèries són resultat del mateix procés, però tenen diferent productor, aquesta relació s'anomenarà una relació de **derivació**. En funció del rol del productor dins el procés, del caràcter de les evidències, i del grau de completesa en la forma com la sèrie documenta totes les accions o tasques del procés generador, una sèrie podrà adquirir el rol de sèrie **primària**, i les altres sèries derivades adquiriran el rol de sèries **secundàries**. La forma de diferenciar aquesta atribució es detalla a la **GDAD**.

Si les dues sèries són resultat de **processos diferents**, cap de les dues sèries adquirirà cap rol específic, més enllà de la constància de l'existència d'una relació. En aquest cas és on es manifestarà directament l'herència del marc de relacions a nivell de procés, i es parlarà de:

- **Sèrie antecedent:** indica que la sèrie origen completa la seva formació abans que s'iniciï la formació de la sèrie destí, i que el contingut informatiu final de la sèrie origen (la matriu de variables) determina el contingut informatiu de les primeres evidències de la sèrie destí. Aquesta relació requereix sempre que els processos generadors de les sèries, en cas de ser diferents, mantinguin també una relació de continuïtat en el mateix sentit que les sèries.
- **Sèrie conseqüent:** indica que la sèrie origen inicia la seva formació un cop s'ha completat la formació de la sèrie destí, i que el contingut informatiu final de la sèrie destí (la matriu de variables) determina el contingut informatiu de les primeres evidències de la sèrie origen. Aquesta relació requereix sempre que els processos generadors de les sèries, en cas de ser diferents, mantinguin també una relació de continuïtat en el mateix sentit que les sèries.
- **Sèrie dependent o relacionada:** indica que la formació de la sèrie origen i destí coincideixen en el temps, però no en les accions o tasques que documenten. Però existeixen tasques d'un procés que tenen una continuació parcial en tasques del procés relacionat, de forma que es poden identificar vincles entre tasques de diferents processos (subprocés o procés compartit), cosa la qual pot generar evidències compartides.

¹⁶⁸ Vegeu també l'apartat 2.3.1.4.2 del capítol 2.

- **Sèrie recapitulativa:** indica que la formació de la sèrie origen és conceptualment posterior a la sèrie destí, malgrat que temporalment puguin ser simultànies en la seva formació. La sèrie origen adquireix la seva condició de sèrie recapitulativa en el moment en què recull una part succinta de les dades d'una o més evidències de la sèrie destí.

Finalment, es poden considerar també l'existència de vincles entre sèries documentals generades per processos diferents i no relacionats funcionalment, sinó per l'objecte o destinatari del procés. En aquest cas les sèries que únicament comparteixen l'objecte mantenen una relació **de contingut**, però no funcional.

Aquestes relacions, llevat de la relació de contingut, queden reflectides a través de la relació **REL012 sèrie documental amb sèrie documental**.

Així doncs, partint de la identificació de la sèrie primària d'un procés, podem sintetitzar aquest mapa de relacions en el següent quadre:

Taula 3-4: Criteris per a l'establiment de relacions entre sèries documentals

Mateix procés?	Coincidència temporal?	Mateix objecte?	Tipus de relació	Caràcter de la sèrie relacionada	Representació de la relació
Sí	Simultània	Sí	Derivació	Secundària	S'informarà l'atribut REL012-001 Rol de la sèrie amb el valor "Relacionada" i l'atribut ENT007-004 Caràcter amb el valor "Secundària".
No	Anterior	Sí	Continuïtat	Antecedent	S'informarà l'atribut REL012-001 Rol de la sèrie amb el valor "Antecedent".
No	Posterior	Sí	Continuïtat	Conseqüent	S'informarà l'atribut REL012-001 Rol de la sèrie amb el valor "Conseqüent".
No	Simultània	No	Dependència	Relacionada	S'informarà l'atribut REL012-001 Rol de la sèrie amb el valor "Relacionada" i l'atribut ENT007-004 Caràcter amb el valor "Primària".
No	Posterior	Sí	Dependència	Recapitulativa	S'informarà l'atribut REL012-001 Rol de la sèrie amb el valor "Recapitulativa".

Mateix procés?	Coincidència temporal?	Mateix objecte?	Tipus de relació	Caràcter de la sèrie relacionada	Representació de la relació
No	Indiferent	Sí	Contingut	Cap	Es derivarà de les instàncies de l'entitat ENT014 Objecte que comparteixin ambdues sèries documentals.

3.2.6.5 Com es relacionen les sèries documentals amb altres entitats

El caràcter interseccional de la sèrie documental es materialitza a través de dues relacions:

- La relació amb el procés que genera la sèrie documental, a través de la relació **REL013 procés amb sèrie documental**.
- La relació amb l'agent que actua com a productor de la sèrie documental, a través de la relació **REL004 agent amb sèrie documental**.

I el caràcter actiu de la sèrie documental es manifesta a través de la relació amb els dos conjunts de regles ja esmentats:

- Les regles per a l'accés i publicitat de la sèrie documental a través de la relació **REL014 sèrie documental amb regla d'accés** i el vincle amb l'entitat **ENT008 Regla d'accés**.
- Les regles per a la conservació, transferència i transformació de les evidències de la sèrie documental a través de la relació **REL015 sèrie documental amb regla de conservació** i el vincle amb l'entitat **ENT009 Regla de conservació**.

3.2.6.6 Com es gestionen les altes, baixes i modificacions de sèries documentals

Les sèries documentals, com a creació operativa que són, es creen i modifiquen en funció de l'evolució en la forma de documentar els processos. Així, canvis en les formes documentals o en els sistemes d'informació, en tant que indicis de canvis en els processos, poden provocar baixes i altes de sèries. Però pel seu caràcter interseccional, i per la dependència que tenen de les dues entitats que interseccionen, els processos i els agents, també es poden veure afectades pels canvis en aquestes entitats.

A part d'aquests canvis, la pròpia revisió del mapa de sèries pot provocar també modificacions en les sèries documentals. De fet, l'optimització del substrat informacional ha de ser l'objectiu principal de qualsevol programa de gestió documental i de la informació, i el mapa de sèries documentals ofereix una base potent per aquesta optimització.

Es detallen a continuació els canvis que poden afectar a les sèries documentals:

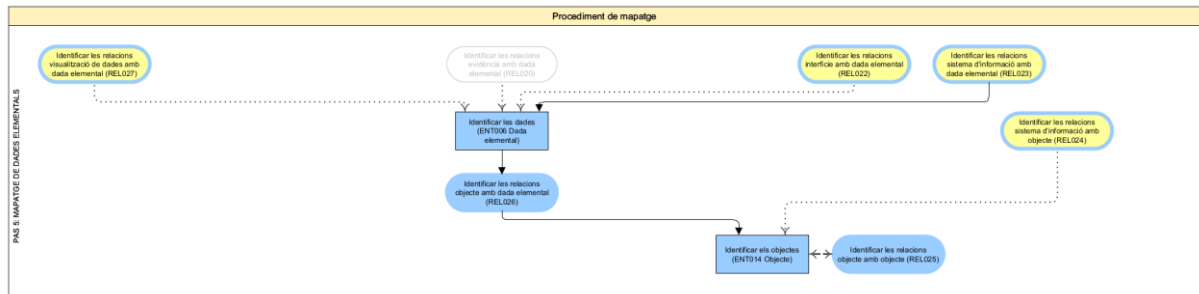
Taula 3-5: Pautes d'actuació per al manteniment del mapa de sèries documentals

ESDEVENIMENT	ACTUACIÓ
Canvis derivats del context orgànic i funcional	
Reestructuració orgànica, creació i supressió d'agents i redistribució de competències	La procedència o productor d'una sèrie documental mai es pot modificar. Per tant, es donarà de baixa la sèrie del productor anterior, informant l'atribut ENT007-009 Data supressió , i es donarà d'alta una nova sèrie documental.
Canvi en les famílies de processos	Els canvis en les famílies de processos no tindran cap afectació sobre les característiques de les sèries.
Canvis que impliquin la baixa i l'alta de processos	El vincle d'una sèrie amb el procés generador no es pot modificar. Per tant, es donarà de baixa la sèrie del procés anterior, informant l'atribut ENT007-009 Data supressió , i es donarà d'alta una nova sèrie documental.
Canvi en el procediment o flux d'execució del procés, incloent l'inici i/o la finalització	Aquests canvis podran implicar modificacions substancials en el mapa de sèries, des de la identificació de noves sèries i el tancament de sèries obsoletes, fins a la modificació de característiques rellevants de les sèries, com el seu caràcter (ENT007-004 Caràcter), la seva granularitat i/o forma d'identificació d'expedients (ENT007-003 Criteri de nominalització), el seu caràcter vital (ENT007-006 Vital) o la quantitat, sentit i tipus de relacions amb altres sèries documentals (relació REL012 i atribut REL012-001 Rol de la sèrie).
Canvis derivats de l'optimització del substrat informacional	
Detecció de diferències en la granularitat de l'expedient (nivell de descripció, o objecte)	Implicarà una homogeneïtzació entre la granularitat de l'expedient i del procés, que es pot reflectir en una modificació del criteri de nominalització (atribut ENT007-003 Criteri de nominalització).

ESDEVENIMENT	ACTUACIÓ
<p>Detecció de denominacions o codis diferents en nominals per a un mateix objecte</p>	<p>Implicarà una homogeneïtzació de les denominacions i codificacions utilitzades pels mateixos objectes, que es traduirà en la modificació del criteri de nominalització en totes les sèries afectades (atribut ENT007-003 Criteri de nominalització).</p>
<p>Existència de sèries secundàries</p>	<p>Per reduir el nombre de sèries secundàries o fer-les irrelevantes i eliminables, caldrà actuar sobre el procés que les genera a través procediment que es descriu a la GDAD. Aquesta operació no ha de provocar necessàriament la baixa de la sèrie documental, però sí modificacions substancials en les seves evidències (entitat ENT004 Evidència documental i sistemes d'informació que les hostatgen), en les característiques de la sèrie (ENT007-004 Caràcter, ENT007-003 Criteri de nominalització, ENT007-006 Vital) i en el seu marc de relacions amb altres sèries documentals (relació RELO12 i atribut RELO12-001 Rol de la sèrie).</p>
<p>Existència d'agrupacions documentals factícies</p>	<p>Implicarà la conversió de les agrupacions en criteris de nominalització, i la descomposició dels expedients factícies en expedients reglats o tècnics, donant de baixa les sèries factícies (atribut ENT007-009 Data supressió) i donant d'alta noves sèries documentals.</p>

3.2.7 Pas 5: mapatge de dades

Figura 3-13: Estructura per al mapatge de les dades



3.2.7.1 Què s'entén per dada

El concepte dada, en el marc d'aquesta guia, es defineix a dos nivells:

- La **dada elemental**. Entenem per dada elemental aquell element d'informació amb unitat de significació pròpia no discursiva. A diferència d'un document o evidència, el qual podem entendre com una oració predicativa (conceptualment sempre té un verb que descriu una acció), una dada elemental no pot representar cap acció, només coses. La dada té doncs una significació atributiva (atribueix una característica a una cosa), no predicativa.
- La **dada instanciada**. La dada elemental és un element propi d'un model lògic. Una mateixa dada elemental pot estar composta de dades de menor dimensió (dades atòmiques). I pot estar-ho en diferents combinacions i amb diferent semàntica. Per exemple, una adreça, entesa com a dada elemental, com estar implementada en un formulari amb un únic camp de text, i en un altre formulari pot aparèixer descomposta en diferents camps de text (carrer, número, pis, escala, etc.). O, per exemple, una magnitud pot estar expressada en un cas en format de text, en un altre de número enter, i en un altre de número amb decimals. La manifestació formal de la dada dependrà de l'entorn (habitualment tecnològic) d'implementació, mentre que la manifestació estructural dependrà de la persona que hagi dissenyat el formulari, model de document o model de dades del sistema d'informació que hi doni suport.

3.2.7.2 Com s'identifica una dada

L'objecte de l'inventari de dades seran les dades elementals que gestioni el conjunt de l'organització. La dada elemental té un component convencional en la seva identificació, atès que és producte d'una abstracció a partir de diferents manifestacions d'aquella dada en documents/evidències, formularis i sistemes d'informació. Per tant, és a partir de la dada instanciada en cada document o sistema d'informació que s'haurà d'obtenir la dada elemental.

L'inventari de dades es construeix de forma progressiva, i s'alimenta a través de dos mètodes:

- A. A través de l'**anàlisi del model de dades intrínsec a cada sistema d'informació** de l'organització. Un buidatge dels models de dades de cada sistema o aplicació ens permetrà obtenir un primer inventari de les dades instanciades en suport informàtic.
- B. A través de l'**anàlisi de les evidències generades per cada procés**. La construcció de la matriu de variables de cada procés donarà com a resultat el diccionari de variables d'aquell procés, el qual, un cop harmonitzat, s'incorpora a l'inventari de dades de l'organització.

El primer mètode permet recollir en el mapa únicament aquelles dades elementals que estan en format estructurat i gestionades mitjançant sistemes d'informació. Ens proporciona per tant un mapa d'aquelles dades elementals que són directament explotables, però no la visió real de totes les dades de l'organització.

Per aconseguir aquesta visió és necessari incorporar també aquelles dades que estan en qualsevol tipus de document-objecte digital, o fins i tot en documents en paper, cosa la qual aconseguirem amb el segon mètode: buidar les evidències de cada procés.

El primer mètode és el que es descriu en aquesta guia. El segon es descriu en la **GDAD**. Detallem a continuació l'aplicació del primer mètode.

3.2.7.2.1 Identificar els objectes descrits

El procés d'identificació de dades elementals candidates dins un conjunt de dades instanciades es basa en la característica principal de la dada elemental, que és l'**objecte que descriu**, atribució que analitzarem amb més detall en la categorització de la dada.¹⁶⁹

¹⁶⁹ Per un desenvolupament d'aquesta idea vegeu l'objectiu 7 a l'apartat 8.1.7 de les conclusions (capítol 8).

En el primer cas (a partir de l'anàlisi del Sistema d'informació) podrem identificar més ràpida i fàcilment la dada, fent un buidatge a partir de models de dades i diccionaris de dades del sistema en qüestió, però difícilment podrem identificar amb prou precisió l'objecte de la dada, ja que l'estarem identificant amb poca informació sobre el seu context funcional. D'aquesta manera, per exemple, un codi de sol·licitud de beca d'un estudiant podrem considerar-lo com una dada descriptiva de l'estudiant, com una dada identificativa d'una beca, o com una dada transaccional corresponent al número d'un expedient de beca. Serà molt difícil determinar una única adscripció objectual de la dada.

Si disposem de la documentació tècnica del sistema d'informació analitzat, i podem accedir al model lògic de dades, podrem utilitzar l'atribució de dades a determinades entitats o classes com a forma d'identificar l'objecte de la dada. Però si només podem accedir al model físic, l'estructura de relacions dels sistema d'informació pot distorsionar aquesta assignació d'objecte.

La forma més precisa, però al mateix temps més laboriosa, de categoritzar les dades en funció de l'objecte que descriuen és a partir de la construcció de la matriu de variables d'un procés. Quan es fa aquesta operació es disposa de la informació sobre el context funcional suficient per assignar amb precisió l'objecte descrit per la dada en primera instància, i per tant poder assignar la dada de forma unívoca a un objecte. Aquesta és l'operació que veurem descrita en la **GDAD**.

Per tant, per identificar possibles objectes descrits per les dades d'un sistema d'informació, farem el següent. Si tenim accés a la documentació tècnica del sistema d'informació, analitzarem el model de dades que es va establir en fase de disseny (preferentment el model lògic). En funció del model tecnològic subjacent, el modelat haurà utilitzat diferents tècniques (modelat relacional, modelat dimensional, modelat orientat a objectes, etc.). A partir d'aquest model, dividirem les entitats del model en dos grups:

- a) Aquelles entitats, classes o dimensions que corresponen a coses existents en el món real.
- b) Aquelles entitats, classes o taules que corresponen a fets o accions.

En aquest procés no tindrem en compte aquelles taules o elements propis únicament del model físic, o creats per exigències de la implementació tecnològica.¹⁷⁰ A partir d'aquesta anàlisi obtindrem un conjunt d'objectes, relacionats funcionalment entre ells i classificats en funció de si són "cosa" o "acció", que incorporarem al mapa com a instàncies de l'entitat **ENT014 Objecte**.

¹⁷⁰ Com és per exemple la resolució de les relacions de grau N en un model relacional.

3.2.7.2.2 Identificar les dades elementals

La següent operació serà identificar les dades elementals. El concepte de dada elemental és una abstracció, que haurem de fer a partir de les dades que apareixen en tota implementació d'un sistema d'informació. Els criteris per a fer aquesta abstracció són els següents.

Agruparem sota una única dada elemental totes aquelles variables que descriuen aspectes o facetes d'un mateix atribut, i que puguin mantenir la significació en una implementació que prengui una forma diferent. Seran dades, doncs, que podran ser objecte d'explotació. Així doncs, en una capçalera de taula que contingui:

Tipus de via	Nom de la via	Número	Pis	Porta
--------------	---------------	--------	-----	-------

Podrem identificar-ho com una única dada elemental, que anomenarem “adreça” o “domicili”. Però si per exemple apareix també el municipi, o les coordenades del domicili, ho podrem identificar com a dades elementals diferents, atès que pot tenir una significació que va més enllà del domicili. Aquesta decisió vindrà determinada en molts casos per la possibilitat d'un posterior tractament estadístic o agregat.

En totes aquelles variables on es representa codificat un valor d'un domini, considerarem que tant el codi com el literal responen a una mateixa dada elemental. Però quan el codi utilitzat no sigui específic d'aquella base de dades o implementació tecnològica, sinó un codi assignat externament al sistema, aleshores s'identificarà com una dada elemental independent (per exemple un DNI/NIF, una matrícula, etc.).

3.2.7.2.3 Harmonitzar les dades

L'harmonització de les dades elementals és un procés continu. Cada cop que s'analitza el model de dades d'un sistema d'informació, o es construeix la matriu de variables d'un procés, s'obté una relació de dades elementals candidates. Aquesta relació s'ha de comparar amb el mapa de dades ja existent. Les dades elementals candidates que no es puguin emparellar amb cap dada elemental ja existent a l'inventari, s'hi incorporaran com a noves dades. Les que s'hi puguin emparellar, confirmaran dades elementals ja existents (i es consignaran com a noves instàncies o aparicions d'aquella dada elemental).

Per facilitar aquesta harmonització, en la denominació de les dades elementals cal utilitzar el nom que millor descriu el significat d'aquella dada, malgrat no coincideixi amb el literal o denominació assignada pel desenvolupador dins del sistema d'informació, atès que una mateixa dada elemental pot aparèixer en columnes amb

noms diferents (per exemple ciutat o municipi). En aquest cas s'optarà per la denominació més precisa o normativa, i així es fixarà.

Les dades elementals ja identificades i harmonitzades generaran noves instàncies de l'entitat **ENT006 Dada elemental**.

3.2.7.3 Com es caracteritzen les dades

Els atributs que utilitzarem per a descriure una dada en l'entitat **ENT006 Dada elemental** són els següents:

- **ENT006-001 Identificador**
- **ENT006-002 Nom**
- **ENT006-003 Tipus de dada**
- **ENT006-004 Descripció**
- **ENT006-005 Regles de validació**
- **ENT006-006 Nivell seguretat RGPD** (específic CEB)
- **ENT006-007 Disponible per interoperabilitat externa** (específic CEB)

El criteri per informar cadascun dels atributs està descrit en el model de dades.

Un atribut que requereix un detall especial és l'atribut **ENT006-003 Tipus de dada**. Es tracta d'un atribut categòric, que no pretén definir la dada en funció del seu tractament informàtic, com es fa en un diccionari de dades, sinó en funció del seu ús i del tipus de significats que pot agrupar. Així doncs, no serà rellevant si en aquella implementació la dada és numèrica, textual o booleana, sinó si es tracta d'una dada identificativa, descriptiva, acadèmica o laboral.

Hi ha una segona entitat que és objecte de caracterització en aquesta fase: l'objecte descrit. Els atributs amb els quals es descriurà l'objecte a l'entitat **ENT014 Objecte** són les següents:

- **ENT014-001 Nom**
- **ENT014-002 Tipus**

L'atribut **ENT014-002 Tipus** és el que serveix per indicar el caràcter referencial o transaccional de l'objecte (cosa o acció).

3.2.7.4 Com es relacionen les dades entre elles

Les dades elementals no mantenen en aquest model cap relació entre elles, llevat de les relacions d'agrupació que es pugui establir a partir dels seus atributs categòrics o de les relacions amb altres entitats.

Però els objectes descrits per les dades sí que poden tenir un nombre elevat de relacions entre ells. En aquesta metodologia no s'explora el desenvolupament ontològic de les relacions entre objectes, pel que fa a la seva condició de coses i accions, i també al context on s'identifiquen.¹⁷¹ Únicament es contempla una relació de comprensió a dos nivells, que s'expressa mitjançant la relació **REL025 objecte amb objecte**:

- **Objecte** pot ser subtipus d'un **objecte**
- **Objecte** pot ser supertipus d'un o més **objectes**

3.2.7.5 Com es relacionen les dades amb altres entitats

La relació principal de les dades és amb els objectes que descriuen. Aquesta vincle queda recollit en la relació **REL026 objecte amb dada elemental**:

- **Dada elemental** descriu un **objecte**

Aquesta relació és l'element bàsic per a la construcció del mapa de dades de l'organització, atès que obliga a identificar les relacions entre els diferents objectes descrits per poder classificar adequadament les dades. Aquesta identificació es pot fer a priori, si es coneix amb suficient detall el model de dades global, o a mesura que es van identificant dades, si aquest model de dades no existeix formalitzat a nivell global o està distorsionat per un creixement no coordinat o poc coherent dels sistemes d'informació de l'organització.¹⁷²

La segona relació rellevant de les dades és amb els sistemes d'informació on apareixen, vincle expressat a través de la relació **REL023 sistema d'informació amb dada elemental**:

- **Dada elemental** pot estar continguda a un o més **sistemes d'informació**

¹⁷¹ Aquest plantejament ontològic s'explora en l'objectiu 7 de les conclusions (apartat 8.1.7 del capítol 8).

¹⁷² Sovint la creació de nous sistemes d'informació en una organització obeeix més a criteris tecnològics, econòmics o d'oportunitat que no pas a criteris de necessitats del model global d'informació.

Les dades elementals poden estar vinculades també a dos elements tecnològics més: les interfícies (**ENT012 Interfície de connexió entre sistemes**) i les visualitzacions de dades (**ENT013 Visualització de dades**). Aquests vincles queden recollits a les relacions **REL022 interfície amb dada elemental** i **REL027 visualització de dades amb dada elemental**:

- **Dada elemental** pot ser traslladada/migrada per una o més **interfícies**
- **Dada elemental** pot aparèixer a una o més **visualitzacions de dades**

Els objectes també mantenen vincles amb elements tecnològics, concretament amb els sistemes d'informació, a través de la relació **REL024 sistema d'informació amb objecte**:

- **Objecte** pot ser unitat de descripció d'un o més **sistemes d'informació**

El vincle de les dades amb les evidències es desenvolupa a la **GDAD**.

3.2.7.6 Com es gestionen les altes, baixes i modificacions de dades

La dada elemental es considera una entitat que, un cop identificada, pràcticament no té variabilitat, més enllà dels ajustaments que es puguin derivar de la detecció de noves ocurrències de la dada que, durant el procés d'harmonització, suggereixin modificar valors dels seus atributs.

L'existència d'una determinada dada en una organització haurà deixat rastre en una sèrie d'evidències, que en alguns casos poden pertànyer a sèries afectades per regles de conservació permanent. Aquesta dada, doncs, seguirà existint en els arxius de l'organització, malgrat no ser present en cap evidència o sistema d'informació vigent. Per aquest motiu aquesta metodologia no es planteja l'eliminació de cap dada un cop incorporada aquesta al mapa de dades, com tampoc s'eliminaran les ocurrències d'aquesta dada en evidències i sistemes d'informació no vigents.

3.3 Guia de dadificació de processos (GDAD)

3.3.1 Estructura i mètode d'aquesta guia

La gestió eficaç d'un procés passa per conèixer i poder millorar el substrat informacional necessari per a gestionar-lo. Entenem substrat informacional com la forma com s'enregistren i documenten les accions d'un procés, i en conseqüència les evidències que es generen amb aquesta acció de documentar.

Tot procés genera una seqüència d'evidències que es creen seguint una lògica procedimental, de manera que una evidència no es pot formalitzar fins que alguna evidència anterior no li aportí les variables d'informació (o dades) que necessita. L'anàlisi i millora d'aquest nexce procedimental és l'objecte d'aquesta guia.

Com s'ha comentat a la introducció de la Guia d'inventari i mapatge (**GMAP**), la metodologia que apliquem consta de dues aproximacions a la realitat informacional d'una organització:

- A. **Enfocament *top-down***. Consisteix en identificar les diferents entitats a partir d'un treball de camp general d'inventari i catalogació, i a continuació construir el mapa de relacions i preparar la base per un detall posterior.
- B. **Enfocament *bottom-up***. Consisteix en treballar a nivell de procés, identificant totes les entitats implicades en un mateix procés, fins al màxim nivell de detall. Aquest treball, repetit per cadascun dels processos, anirà alimentant el mapa global de l'organització.

En aquesta guia es prioritza l'estratègia *bottom-up*: s'aprofundeix en un procés fins arribar al nivell de dada, per a continuació enllaçar aquest procés, a través de les dades, amb la resta de processos de l'organització. Però per a fer possible aquest enllaç, cal prèviament descompondre l'estructura del procés en les unitats mínimes que el componen, les dades, per tal de reconstruir-lo posteriorment en una forma intrínsecament vinculada al conjunt d'informació de l'organització, i als sistemes d'informació que gestionen aquesta informació. És el procés que anomenem "dadificació": transformació d'un procés basat en documents (evidències-objecte),¹⁷³ en un procés basat en dades.

No és prudent transformar un procés amb disfuncions. Per tant, prèviament a la dadificació cal depurar l'estructura del procés, materialitzada en la seqüència d'evidències que genera, i en com s'agrupen aquestes en sèries documentals. S'apliquen doncs en aquesta operació els principis i pautes enumerats en la **GMAP**, per

¹⁷³ En aquesta guia utilitzarem com a sinònims els termes "document-objecte" i "evidència-objecte". Per a una definició completa del concepte vegeu l'estat de la qüestió.

reduir el substrat informacional del procés a una sèrie documental única, que és la que posteriorment es dadificarà.

Es treballa, doncs, als dos nivells que proposa la metodologia APS:¹⁷⁴

1. El primer nivell (optimitzador) parteix d'una lectura focal del procés, com a ens independent dins del conjunt de processos de l'organització, i es concentra en la millora del procés objecte d'anàlisi. L'objectiu és millorar el desenvolupament del procés sense fer canvis substancials en la naturalesa del seu substrat informacional.
2. El segon nivell (transformador) amplia l'abast de les modificacions, i preveu fer canvis substancials en la naturalesa del substrat informacional que permetin millorar la presa de decisions. En aquest cas l'objectiu és transformar des de la base la forma com es documenta el procés, i fer-ho de forma integrada amb el mapa informacional de l'organització.

Aquest conjunt d'operacions complementen doncs la metodologia exposada en la **GMAP**, amb la qual aquesta guia comparteix conceptes, operativa i model de dades. En el següent gràfic s'indiquen les parts de la metodologia que s'activen amb l'operativa de la **GDAD**.

¹⁷⁴ Per a la metodologia APS vegeu l'apartat 3.1.2.4 d'aquest capítol.

3.3.2 Pas 1: descripció

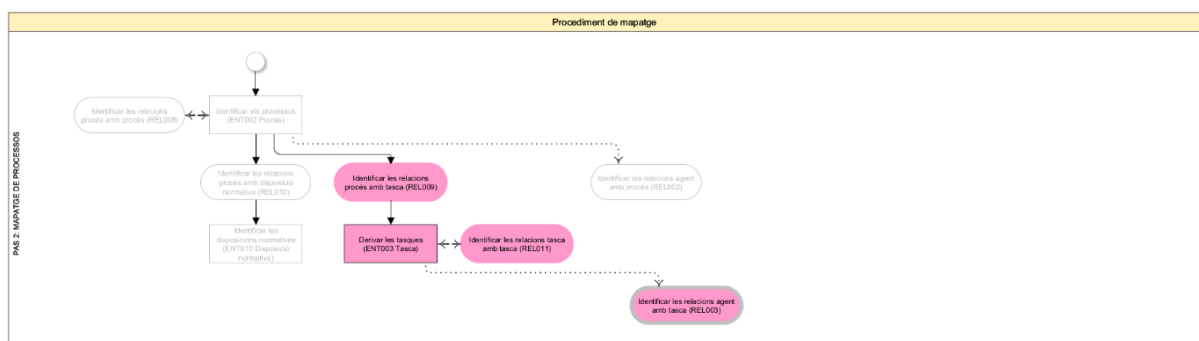
El primer pas és la descripció detallada del desenvolupament del procés, és a dir del procediment. El procés ha estat prèviament identificat mitjançant la Guia d'inventari i mapatge (**GMAP**), i n'han estat determinats clarament els seus límits, és a dir l'acció que l'inicia i el finalitza. Per tant, es parteix d'una primera descripció estructurada del procés, el model de descripció de la qual és el que s'indica a l'entitat **ENT002 Procés** del model de dades.

La categorització d'un procés implica també la descomposició del procés en tots els elements que el componen. Aquests elements, a part dels agents que hi intervenen, són les tasques que realitza cada agent dins del procés, les evidències que rep o genera l'execució de cada tasca, i els elements que es relacionen amb cada evidència, com són els models de document per a les evidències tipificades, les dades que componen cada evidència i els sistemes d'informació que les contenen.

La descripció del desenvolupament del procés, doncs, es construeix en primer lloc a partir de la seva descomposició en elements funcionals de menor dimensió, o tasques (**ENT003 Tasca**), tal com s'indica a continuació.

3.3.2.1 Descriure i, si escau, diagramar el procés

Figura 3-15: Estructura per a la descripció i diagramació del procés



L'estructuració del desenvolupament del procés comença a partir de la seva descripció. És important tenir present que la metodologia no implica ni requereix cap representació gràfica del procés, tot i que és recomanable treballar-hi. El que sí que requereix la metodologia és una estructuració de les dades del procés, d'acord amb el model de dades descrit al capítol 4.¹⁷⁵

¹⁷⁵ Segons el concepte de “categorització per descomposició” esmentat en l'apartat 3.2.4.3.2.

És difícil obtenir una descripció estructurada directament a partir del treball de camp, el qual es basa habitualment en entrevistes, qüestionaris, observació directa i anàlisi documental, atès que les fonts són diverses i poden, en molts casos, contenir informació contradictòria. Per aquest motiu, el primer que cal fer és establir l'ordre de buidatge de les fonts, en funció de la seva fiabilitat estimada, i en cas que aquesta no es pugui fixar, prioritant sempre les fonts formals (normativa, procediments de treball, etc.) davant de les fonts menys formals (entrevistes, observacions, etc.).¹⁷⁶

Un cop recollida la informació, s'anirà marcant dins de cada font les mencions a dues entitats:

- Els agents
- Les tasques que realitza cada agent

Els agents humans sempre intervindran en el procés en forma de rols (unitats administratives, càrrecs, etc.) o ens impersonals, mai amb noms de persona. Quan els agents siguin automatitzats, i no siguin clarament equiparables a un rol o càrrec, es farà una imputació de les tasques d'aquest agent en el procés a la unitat administrativa a qui correspondria la competència o responsabilitat sobre la tasca efectuada, malgrat aquesta unitat no sigui la propietària o promotora de l'agent automatitzat.

Malgrat s'hagi definit prèviament un mapa d'agents, és possible que durant la descripció del procés es puguin identificar nous agents no detectats anteriorment, especialment agents externs. En aquest cas, la pauta d'identificació dels agents i el seu model de descripció és el que s'indica a l'entitat **ENT001 Agent** del model de dades.

3.3.2.1.1 Identificació de les tasques

La identificació de tasques dins el procés és un repte important. La mateixa dificultat de diferenciar límits en seqüències d'accions per tal de separar un procés d'un altre es planteja en el cas de les tasques. I, de la mateixa manera, el nivell de detall en la identificació de tasques també es desenvolupa per nivells de concreció, per la qual cosa acaba adquirint un sentit més convencional que científic.

Les tasques s'identificaran bàsicament a partir de la seva relació amb els agents que les executen:

- A. Cada cop que un agent passi l'execució del procés a un altre agent, es finalitzarà una tasca del primer agent i s'obrirà una nova tasca del segon agent.

¹⁷⁶ Per un major detall sobre la recollida i buidatge de les fonts d'informació vegeu l'apartat 3.4 d'aquest capítol.

- B. Quan un mateix agent executi una part important del procés, es diferenciarien tantes tasques d'aquest agent com productes o resultats concrets (productes parcials del procés) generi aquest agent.

Tot i així, sovint poden aparèixer dubtes en aquesta identificació. Per aquest motiu es detallen a continuació algunes consideracions que poden ajudar en aquesta identificació:

- Les tasques són únicament aquelles accions realitzades de forma voluntària per un agent (humà o màquina).
- El que no depèn per a la seva execució d'un agent implicat en el procés es tracta com a esdeveniment, no com a tasca¹⁷⁷. Això no obsta per a que, als efectes de descripció, s'assigni l'esdeveniment a un agent. Els esdeveniments poden ser de dos tipus:
 - Basats en calendari (programats o programables)
 - Imprevisibles (però detectables)
- El flux del procés pot contenir també bifurcacions. Aquestes bifurcacions poden estar guiades per decisions o accions d'un agent (tasques), o per esdeveniments.
- En principi, considerarem que es tracta d'**una sola tasca** quan es donin les següents condicions:
 - Que una seqüència d'accions la realitzi un sol agent
 - Que les accions es puguin realitzar de forma contínua i sense pausa
 - Que les accions generin en acabar la seqüència una o varies evidències relacionades
- Els indicadors que denoten que cal partir una seqüència d'accions en **més d'una tasca** són:
 - Que les accions estan realitzades per més d'un agent
 - Que s'hi intercalen temps d'espera (d'esdeveniments) o pauses de forma natural
 - Que s'hi intercalen decisions que poden bifurcar el flux del procés
 - Que s'hi intercalen accions de tramesa o comunicació d'evidències

¹⁷⁷ Hi ha qui contempla i representa com a esdeveniments també aquelles tasques realitzades per un agent extern a l'organització, especialment quan el procés no considera possible actuar sobre les accions d'aquest agent extern.

El flux del procés vindrà determinat per la relació de cada tasca amb la seva o seves tasques anteriors. La tasca inicial serà aquella que no tingui antecedent, i la tasca final aquella que no sigui antecedent de cap altra tasca. La disposició en seqüència d'aquestes tasques dependrà de la forma del procés. Els processos, amb caràcter general, poden prendre dues formes:¹⁷⁸

- A. Forma preferentment **seqüencial** (les tasques es desenvolupen de forma consecutiva). És el disseny típic dels processos que s'anomenen **reglats**, els quals habitualment tenen una definició de la seva execució recollida en alguna norma (procediment).
- B. Forma preferentment **radial** (les tasques es desenvolupen sense un ordre o seqüència predefinits, i tornen a un agent que fa la funció de distribuïdor). És el disseny habitual dels processos que s'anomenen **no reglats**, els quals no tenen una definició normativitzada del seu desenvolupament; el seu desenvolupament no està necessàriament predefinit, sinó que pot respondre a múltiples combinacions de seqüència, i sovint orbiten al voltant d'un agent que concentra el pes del procés i fa les vegades de distribuïdor i impulsor de les diferents tasques.

Utilitzant doncs tasques, esdeveniments i bifurcacions es representarà la seqüència del procés.

3.3.2.1.2 Descripció de les tasques

A mesura que es vagin identificant tasques, s'aniran informant les seves descripcions. Aquestes descripcions corresponen a l'entitat **ENT003 Tasca** del model de dades, i es componen dels següents atributs:

- **ENT003-001 Identificador**
- **ENT003-002 Nom**
- **ENT003-003 Descripció**
- **ENT003-004 Tipus executiu**
- **ENT003-005 Tipus en el flux**
- **ENT003-006 Condició de bifurcació**
- **ENT003-007 Data creació**
- **ENT003-008 Data supressió**

¹⁷⁸ Sobre les formes procedimentals dels processos vegeu l'apartat 2.3.2.2 del capítol 2.

En la descripció de les tasques, cada tasca s'assignarà a un sol agent (a través de la relació **REL003 agent amb tasca**), i també al procés al qual pertany (a través de la relació **REL009 procés amb tasca**).

A continuació s'indicarà quina és la tasca antecedent i quina la conseqüent (utilitzant l'atribut **REL011-001 Rol de la tasca** de la relació **REL011 tasca amb tasca**). Com s'ha comentat anteriorment, en cas que una tasca tingui més d'un agent, es crearan tantes tasques com agents hi intervinguin, o bé se li assignarà un agent col·lectiu.

Les tasques poden tenir diverses tasques anteriors, i diverses de conseqüents; en aquests casos, s'utilitzaran bifurcacions (*gateways*) per indicar que existeix una coincidència (en el cas de les anteriors) o divergència (en el cas de les conseqüents) en la conseqüència de les tasques. La bifurcació s'assignarà a l'agent que intervingui en determinar la direcció del flux (si és una decisió), o a l'agent que hagi realitzat la tasca immediatament anterior. El sentit de la bifurcació s'informarà en l'atribut **ENT003-006 Condió de bifurcació**.

A part dels *gateways*, les tasques també poden utilitzar altres tipus de connectors, que permeten representar els esdeveniments. Sempre s'iniciarà la descripció amb un connector d'inici de procés, i es finalitzarà amb un connector de fi de procés. Quan dins el procés es produeixi un vincle amb un altre procés, o bé s'espera un esdeveniment extern, resultat d'un altre procés, s'indicarà amb un connector de relació. El tipus de connector s'informarà en l'atribut **ENT003-005 Tipus en el flux**.

En funció del tipus d'esdeveniment, s'assignarà a un agent o un altre:

- a) Un esdeveniment d'inici de procés s'assignarà a l'agent que executi la primera tasca del procés.
- b) Un esdeveniment de fi de procés s'assignarà a l'agent que hagi realitzat la darrera tasca del procés.
- c) Un esdeveniment de comunicació o missatge s'assignarà a l'agent que rebi la comunicació.
- d) Un esdeveniment de calendari o termini s'assignarà a l'agent que hagi de realitzar la següent tasca un cop finalitzi el termini.
- e) Un esdeveniment de vincle amb un altre procés (subprocés), o bé d'interrupció del procés, s'assignarà a l'agent que hagi realitzat la tasca immediatament anterior.

La descripció de les tasques pot incloure també elements de caracterització. Es poden utilitzar elements d'informació (escrita o visual) per indicar que una tasca és plenament manual, plenament automatitzada (desatesa), o que implica la interacció d'un humà amb una màquina. El caràcter humà o automatitzat de la tasca s'informarà en l'atribut **ENT003-004 Tipus executiu**.

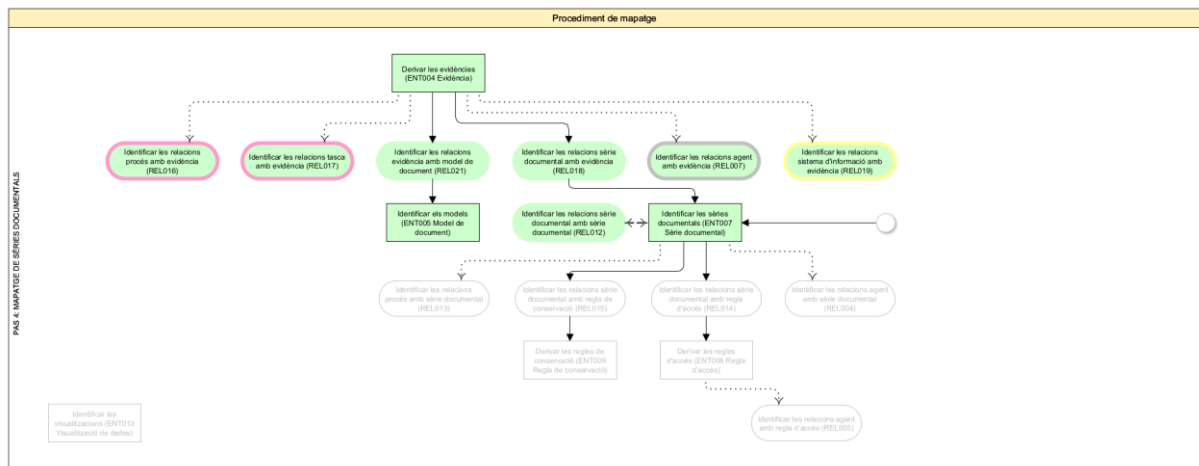
Un cop es disposa d'aquesta descripció estructurada del procés, és possible representar-la de forma gràfica. Existeixen nombroses notacions per a representar un flux de tasques, que permeten incloure també nombrosos elements visuals per a la caracterització de les tasques. Malgrat no és preceptiu elaborar la representació gràfica, i en processos molt llargs pot esdevenir de difícil lectura, habitualment és un recurs útil per contrastar amb els agents l'encert de la definició, i esmenar-ne els errors mitjançant revisions iteratives.

A partir de la identificació i descripció de les tasques podem arribar a identificar també nous vincles entre processos, o fins i tot replantejar-nos la identificació inicial dels processos en la fase de mapatge:

- La relació de cada tasca amb la seva o seves tasques anteriors ha de ser sempre una relació necessària. És a dir, sempre s'ha d'acabar executant el procés sencer, seguint algun dels flux possibles d'acord amb les bifurcacions previstes, i sense solució de continuïtat. Si en algun cas un procés pot interrompre el seu flux d'execució sense estar previst en una bifurcació, cal valorar si cal identificar dos processos relacionats enlloc d'un sol procés.
- L'altre element que ajuda a la identificació són els sovint anomenats "subprocessos", és a dir conjunts de tasques que s'executen en un procés però formen part d'un altre procés. En aquest cas, enlloc d'utilitzar el concepte de subprocés utilitzarem una crida a un altre procés, amb retorn al procés origen o finalització en el procés cridat, i considerarem ambdós processos relacionats al mateix nivell o per una relació de transversalitat o consecutivitat, segons si el flux retorna al procés origen o finalitza en el procés cridat.

3.3.2.2 Derivar el circuit documental del procés

Figura 3-16: Estructura per a la derivació del circuit documental del procés



El següent pas consisteix a identificar, a partir de la descripció del flux del procés, la relació d'evidències que genera aquest procés. El conjunt d'unitats d'informació rebudes o generades durant l'execució d'un procés pels agents que hi participen, les quals esdevenen prova o evidència de com s'ha executat cada tasca o acció del procés, rep habitualment el nom de "circuit documental". En aquesta metodologia, pel fet de parlar d'evidències, també utilitzarem el terme "circuit evidencial".¹⁷⁹

Cada tasca del procés necessita una informació per executar-se, i en portar-se a terme genera una nova informació. La informació que necessita per executar-se prové del que anomenem evidències d'entrada de la tasca, i la informació que genera es condensa en evidències de sortida. Així, cada tasca pot rebre una o més evidències d'entrada, i generar una o més evidències de sortida. La possible repetició d'informació no és rellevant en aquesta fase de l'anàlisi, on només descriurem el procés tal com es porta a terme.

Tradicionalment el conjunt d'evidències tipus o circuit documental d'un procés s'ha obtingut agafant un o més expedients d'aquell procés i analitzant el seu contingut documental. Però aquest sistema no és eficaç en entorns de gestió parcialment dadificats, on una part important de les evidències rauen captives de diferents sistemes d'informació, i no ho fan en format de document-objecte sinó de dades estructurades. Per aquest motiu, es proposa a continuació un procediment diferent:

¹⁷⁹ Per a més informació sobre la significació i equivalències del concepte d'evidència en aquesta metodologia, vegeu l'**ENT004 Evidència** del model de dades.

3.3.2.2.1 Identificació de les evidències

Per extreure les evidències d'informació d'un procés (relació **REL016 procés amb evidència**), es farà la pregunta, per cada tasca, i si és possible seguint la seva seqüència d'execució, de quines evidències rep i utilitza per al seu desenvolupament (evidències d'entrada), i on i en quina forma deixa enregistrada la nova informació que genera (evidències de sortida). Les formes de recollida de la informació seran les mateixes que per a les tasques del procés: un treball de camp basat en entrevistes, observació directa d'aplicacions informàtiques i obtenció d'exemplars de documents.

Les pautes per identificar les evidències seran les següents:

Per als **documents-objecte**:

- Quan una tasca generi un document formalitzat, amb una forma fixa i un contingut estable, aquest document es considerarà una evidència tipus "Document". Això serà així amb independència del suport (paper o digital), del lloc d'emmagatzemament (xarxa local, gestor documental, etc.) i de la forma de generació (automatitzada, manual, etc.).
- Cada manifestació d'una evidència generada a partir d'un document-objecte original es considerarà una nova evidència. En aquest cas, s'identificarà amb el mateix codi i nom que el document original, però afegint un sufix al codi (a, b, c...), i indicant en la descripció el tipus de relació amb el document original (còpia simple o còpia autèntica).
- Quan una tasca generi qualsevol tipus de comunicació o missatge, malgrat que no tingui una forma fixa i un contingut estable, s'identificarà també com una evidència del tipus "Document", i a manca de formalització específica es caracteritzarà com a correu electrònic.
- Qualsevol document-objecte que sigui tramès generarà un mínim de dues evidències en el circuit documental del procés: una per al destinatari, i una altra per al remitent. Ambdues tindran el mateix codi i nom d'evidència, però afegint novament un sufix al codi (a, b, c...).

Per als **sistemes d'informació**:

- Quan una tasca impliqui la consulta de dades en pantalla per part d'un usuari, i/o informar noves dades a través d'una pantalla d'entrada de dades o formulari, es donarà d'alta una evidència corresponent a aquesta interacció. Cada interacció amb un determinat sistema realitzada durant una tasca donarà origen a una evidència.
- Si la interacció és amb diferents sistemes, es donarà d'alta una evidència per cada interacció i sistema.

- Al llarg d'un procés, poden haver-hi múltiples interaccions amb un mateix sistema. Cada interacció donarà origen a una nova evidència.
- Les interaccions automatitzades de generació o transferència de dades també es donaran d'alta com a sistemes, quan formin part de l'execució d'una tasca i hagin estat atribuïdes a un agent.

El resultat d'aquesta identificació serà la creació d'una sèrie d'instàncies de l'entitat **ENT004 Evidència**, i la vinculació de cada evidència a les corresponents tasques, en qualitat d'entrada o sortida de la tasca, a través de l'atribut **REL017-001 Sentit de l'evidència** de la relació **REL017 tasca amb evidència**.

3.3.2.2 Descripció de les evidències

La caracterització de cada evidència inclou els elements descriptius que es detallen a l'entitat **ENT004 Evidència** del model de dades, i que són els següents:

- **ENT004-001 Identificador**
- **ENT004-002 Nom**
- **ENT004-003 Tipus**
- **ENT004-004 Caràcter**
- **ENT004-005 Format**
- **ENT004-006 Autenticació**
- **ENT004-007 Registre d'entrada**
- **ENT004-008 Registre de sortida**
- **ENT004-009 Essencial**
- **ENT002-010 Data creació**
- **ENT002-011 Data supressió**

Els principals elements a tenir en compte en aquesta caracterització són els següents:

- **La codificació de les evidències.** Una mateixa evidència pot ser replicada diverses vegades en el curs del mateix procés, de forma més o menys fidedigna. És a dir, podem trobar múltiples còpies o exemplars d'una mateixa evidència (des d'un punt de vista formal i de contingut informatiu) dins un mateix procés. Atès que l'eliminació de les duplicitats és un dels punts rellevants del procés de dadificació, és important recollir aquest vincle de contingut entre diferents evidències del procés. Per aquest motiu, es recomana assignar un número únic

a cada evidència del procés (a l'atribut **ENT004-001 Identificador**), llevat d'aquells casos en què es tracta de rèpliques de la mateixa evidència, cas en el qual es mantindrà el mateix número i es complementarà amb un sufix (una lletra per exemple), la qual identificarà l'exemplar concret d'aquella evidència.

- **El tipus d'evidència**, que consignarem en l'atribut **ENT004-003 Tipus**. Una evidència pot aparèixer en el procés en forma de document-objecte, o bé en forma d'agregat de dades dins un sistema d'informació. En el primer cas, la ubicació de l'evidència no és rellevant, atès que es tracta d'un objecte que seguirà existint com a tal amb independència de l'entorn on estigui hostatjat. Però en el segon cas, l'evidència és captiva del sistema que la conté, i no se'n pot desvincular si no és generant una nova evidència en forma d'objecte. Per tant, en el cas de les evidències basades en dades, caldrà indicar el sistema que conté l'evidència, utilitzant la relació **REL019 sistema d'informació amb evidència**. Però, atès que un mateix sistema pot contenir moltes més dades de les que es poden utilitzar en un procés en concret, és important indicar quines són les dades del sistema que participen d'una tasca concreta d'un procés. I això es farà creant una evidència de tipus "Dades", que equival, en el cas de les evidències d'entrada, a una pantalla de visualització de dades d'un sistema, i en el cas de les evidències de sortida, a una pantalla d'entrada de dades a un sistema.
- **El caràcter de l'evidència**, que consignarem a l'atribut **ENT004-004 Caràcter**. El procés de dadificació tindrà en compte només les evidències autèntiques, les que tenen presumpció de veracitat. Per aquest motiu, és cabdal determinar l'autenticitat de les evidències, el seu caràcter equivalent a original. L'existència de sistemes d'autenticació (signatures, segellats, autenticacions algorítmiques o pistes d'auditoria), que informarem en l'atribut **ENT004-006 Autenticació**, determinarà aquest valor com a condició complementària, però no exclusiva. La determinació exclusiva del caràcter d'original d'una evidència li atorgarà el fet de ser la primera evidència en aquella forma (el primer exemplar) generat per una tasca del procés, malgrat encara no hagi estat autenticada amb cap mecanisme fefaent.
- **La destinació de l'evidència**. Tota evidència té un agent productor, i pot ser usada en el curs del procés per diversos agents destinataris. El pas d'una evidència del productor fins al seu destinatari final, a través de N destinataris intermedis, s'anomena circuit documental. Als efectes del procés de dadificació, només és realment rellevant el destinatari final de l'evidència, aquell que en serà el custodi definitiu un cop finalitzat el procés. La representació dels destinataris intermedis pot ajudar en la simplificació i optimització del procés, però no és determinant en la seva anàlisi documental. Per aquest motiu, l'atribut **REL007-001 Rol de l'agent** de la relació **REL007 agent amb evidència** només contempla dos possibles valors: creador de l'evidència, i destinatari final. Si

l'evidència acaba emmagatzemada en un sistema d'informació, l'agent responsable d'aquell sistema, o de les funcions que automatitza aquell sistema, serà considerat el custodi de l'evidència.

- **El model de document.** Moltes evidències poden tenir formalitzada la seva forma, estructura i contingut, sigui en forma de plantilles, models de redactat, formularis, etc. En aquests casos, és important indicar també el model de document aplicat a aquella evidència, el qual en determinarà aspectes rellevants pel que fa al contingut informatiu i autenticitat. Aquest vincle s'expressarà a través de la relació **REL021 evidència amb model de document** amb la creació d'una instància de l'entitat **ENT005 Model de document**, la qual es descriurà amb els següents atributs:

- **ENT005-001 Identificador**
- **ENT005-002 Versió**
- **ENT005-003 Nom**
- **ENT005-004 Tipus**
- **ENT005-005 Format**
- **ENT005-006 Data creació**
- **ENT005-007 Data supressió**

El resultat d'aquesta fase serà una relació caracteritzada de totes les evidències d'informació que genera el procés, i que en conseqüència en documenten el desenvolupament. Sobre aquesta relació es treballarà durant la resta del procés de dadificació.

3.3.2.3 Identificar les tasques substantives del procés

Totes les tasques del procés no tenen el mateix valor. Hi ha tasques que es poden suprimir sense que aquesta supressió afecti la naturalesa del resultat del procés. Però n'hi ha d'altres que, en cas de supressió, afecten de forma crítica el resultat del procés, o fins i tot poden impedir la continuació de l'execució del procés. Aquest caràcter va més enllà del fet que se situïn en el camí crític del procés, atès que el camí crític pot contenir diferents tipus de tasques de diferent rellevància. Aquestes tasques, les que afecten substancialment el resultat del procés, s'anomenen tasques substantives. Un exemple de tasques substantives en el procés de concessió d'una beca serien la sol·licitud de beca, la decisió sobre la concessió (informe vinculant, per exemple), i la resolució de

concessió. Tasques no substantives serien les tasques de tramesa de l'informe a la unitat que ha de resoldre, l'emissió d'un informe no vinculant, etc.

En un procés correctament dissenyat, tota tasca del procés ha de quedar documentada, és a dir, ha de generar evidències d'informació suficients per a que hom pugui conèixer com va esdevenir aquella tasca. Però l'essència del discurs d'un procés, la seva síntesi informativa, es concentra en les accions substantives. És a dir, explicar les accions substantives és sempre suficient per explicar la globalitat del procés. Per aquest motiu, des del punt de vista informatiu només són rellevants les evidències generades per les tasques substantives del procés.

Per tant, un cop identificades les tasques substantives del procés, totes aquelles evidències generades per aquestes tasques es caracteritzaran com a evidències "essencials" en l'atribut **ENT004-009 Essencial** de l'entitat **ENT004 Evidència**.

3.3.2.4 Agrupar en sèries documentals les evidències del procés

El circuit evidencial d'un procés en proporciona un marc objectiu per identificar totes les sèries documentals que genera un procés, definir-ne les característiques, i relacionar-les entre elles. I ens permet fer-ho amb molta més precisió que si apliquem l'enfocament *top-down*, descrit a la **GMAP**. En conseqüència, sempre que sigui possible és recomanable identificar les sèries documentals per aquest mètode, de forma posterior a identificar les evidències, i no pas directament a partir del procés.

Recuperant el caràcter interseccional de la sèrie documental, inferirem que cada procés generarà una sèrie documental per cada agent que hi participi activament, generant algun tipus d'evidència. Per tant, un procés amb quatre agents participants automàticament ens generarà quatre sèries. Però hi ha un segon mètode més objectiu per a fer-ho, directament a partir de les evidències.

Per derivar d'un procés les sèries documentals que genera, tindrem en compte únicament el destinatari final de les evidències. Així doncs, ordenant i agrupant la relació d'evidències pel destinatari final obtindrem el número de sèries que genera un procés, coneixerem els seus productors (l'agent destinatari) i la seva composició documental (relació ordenada procedimentalment de les evidències d'informació de la sèrie). Amb aquesta informació, ja podrem instanciar les sèries, en cas que encara no existeixin, a l'entitat **ENT007 Sèrie documental**, i vincular-les a cada evidència a través de la relació **REL018 sèrie documental amb evidència**. Per a la caracterització de les sèries utilitzarem les indicacions que es donen en l'apartat "Pas 4: mapatge de les sèries documentals" de la **GMAP** (apartat 3.2.6), llevat del caràcter i de les relacions entre sèries, que es definiran més endavant.

3.3.2.5 Identificar els agents promotors o competents del procés

En un procés participen diversos agents, que poden pertànyer a una sola o a diverses organitzacions. La vinculació de tots els agents amb el procés no té la mateixa intensitat; hi ha agents que tenen un paper més actiu en el procés, sigui perquè n'executen una part significativa de tasques, sigui perquè en decideixen elements determinants del resultat. D'altres agents poden tenir una vinculació menys intensa, sovint de caràcter complementari, que pot prendre la forma d'assessorament o de diferents formes de suport. En el primer cas, parlarem d'agents competents en el procés (o promotors del procés), i en el segon parlarem d'agents complementaris.

Per tant, per tal de determinar quins són els agents promotors o competents del procés (habitualment un sol agent per organització i procés), haurem d'analitzar:

1. **La quantitat i caràcter de tasques que executa cada agent.** A major nombre de tasques, major competència sobre el procés, que s'incrementarà si les tasques executades són preferentment substantives.
2. **El flux de les tasques.** En un procés seqüencial, la presència recurrent d'un agent en el camí crític del procés li atorgarà candidatura d'agent competent. En un procés radial, la presència d'un agent al vèrtex majoritàriament disparador de tasques cap a altres agents el farà candidat a agent promotor o competent.
3. **El caràcter decisiu de les tasques.** La capacitat de decisió executiva d'un agent, especialment en aquells punts del procés on hi ha bifurcacions (*gateways*) contribuirà, juntament amb els dos factors anteriorment esmentats, a adjudicar-li la condició d'agent promotor o competent.

3.3.2.6 Caracteritzar i relacionar les sèries documentals del procés

El darrer pas de la descripció del procés és la caracterització de les sèries documentals identificades pel que fa a la seva tipologia. D'acord amb això, declararem:

- A. Com a **sèries primàries**, aquelles sèries generades dins de cada organització per agents considerats promotors o competents del procés analitzat.
- B. Com a **sèries secundàries**, aquelles sèries generades dins de cada organització per agents no considerats promotors ni competents del procés analitzat.

Deixarem constància d'aquesta atribució a l'atribut **ENT007-004 Caràcter** de cada sèrie, amb el valor corresponent (primària o secundària), i també vinculant les sèries derivades amb el valor corresponent en l'atribut **REL012-001 Rol de la sèrie** de la relació **REL012 sèrie documental amb sèrie documental**.

Quan dins una mateixa organització, per a un mateix procés, identifiquem dos o més agents competents, i per tant dos o més sèries primàries, les considerarem també com a **sèries relacionades** entre elles (atribut **REL012-001 Rol de la sèrie** de la relació **REL012 sèrie documental amb sèrie documental**), llevat que cobreixin seqüències de tasques consecutives dins el procés, l'equivalent de subprocessos, cas en el qual les considerarem com a part d'una relació de continuïtat (antecedent/conseqüent).

En tots els casos, el procés serà el mateix i l'agent diferent per a cada sèrie documental.

3.3.3 Pas 2: optimització

Quan parlem d'optimització, ens referim a com podem aconseguir, mitjançant la reenginyeria d'un procés, que aquest generi el resultat documental òptim des del punt de vista de l'arxivística i la gestió documental. Per tant, el primer que cal fer és objectivar quin i com és aquest resultat òptim.

Les evidències d'informació generades per un procés generen informació redundant. Les mateixes variables d'informació apareixen en diferents evidències, amb diferent grau d'autenticitat. El procés d'optimització documental s'adreça, doncs, a aconseguir que el procés en conjunt es documenti amb el mínim de redundància de les variables, i amb el major grau d'autenticitat possible. En aquest punt de la metodologia, on encara no s'està treballant al nivell de variable d'informació, això implica treballar sobre aquelles evidències a les quals es pot estimar un major grau d'autenticitat i una major completesa en el nombre de variables d'informació.

El resultat documental òptim d'un procés és aquell que correspon, en l'àmbit de les tecnologies de la informació, a l'ideal de la dada única.¹⁸⁰ Un procés estarà documentat de forma òptima:

- Quan tingui un únic exemplar de cada variable d'informació que han generat (creat o utilitzat) els agents del procés, amb informació veraç i autèntica.

Atès que això només és possible quan el procés està íntegrament dadificat, cosa la qual veurem més endavant, en els processos basats en evidències-objecte la condició s'adequa al següent requisit:

- Quan tingui una única evidència d'informació per a cada acció o tasca del procés (o les que es derivin de forma natural de l'execució de la tasca), amb caràcter d'original i, per tant amb presumpció de veracitat i autenticitat.

És a dir, el resultat documental òptim implicaria una correspondència 1 a 1 entre el procés i la sèrie documental (procés de sèrie única). Però això tampoc és senzill d'aconseguir, perquè la intervenció de diferents agents en el procés introdueix un component orgànic que intersecciona amb el component funcional en base al qual es feia la identificació sèrie-procés, i també perquè la participació de certs agents en el procés, especialment quan pertanyen a organitzacions diferents, és sempre competencialment rellevant. Per aquest motiu, no es treballa a nivell de procés ni d'agent, sinó de sèrie documental, i es determina que la documentació serà òptima:

- Quan cada organització que participa en un procés conservi, en acabar el procés, una única sèrie primària (llevat d'aquells casos en què la competència és compartida, i hem identificat sèries primàries relacionades dins el mateix

¹⁸⁰ Vegeu l'estat de la qüestió per al detall i evolució del concepte de dada única.

procés i organització), que documenti totes les tasques substantives del procés, i que ho faci amb evidències d'informació originals (amb presumpció de veracitat i autenticitat).

Per tant, l'optimització documental consistirà en aconseguir, mitjançant la reenginyeria del procediment, que el procés quedi documentat complint aquesta condició.

3.3.3.1 Analitzar les evidències essencials de les sèries primàries

Un cop identificades i caracteritzades les sèries, el següent pas consisteix a verificar si la sèrie identificada com a primària dins cada organització compleix els requisits de:

- **Completesa:** verificar si totes les tasques substantives del procés tenen una evidència d'informació que les descriu dins l'expedient tipus d'aquesta sèrie documental.
- **Originalitat:** verificar si les evidències d'informació de la sèrie primària categoritzades com "essencials" tenen la condició d'originals o còpies autèntiques.

El més habitual és que el resultat no sigui correcte, i calgui fer modificacions en els circuits documentals de les evidències per rectificar la composició del resultat documental.

3.3.3.2 Modificar el procés per a que generi el resultat documental òptim

L'optimització documental no sempre requereix modificacions dràstiques del flux del procés. En funció del tipus de disfunció, les opcions de modificació poden ser diferents. En la taula següent s'indiquen els defectes més habituals i les possibles opcions de modificació.

Figura 3-17: Pautes d'actuació per optimitzar el resultat documental del procés

Defecte del resultat documental	Possible anomalia procedimental causant	Solució procedimental
Alguna tasca substantiva no està documentada en la sèrie primària.	La tasca no està generant cap tipus d'evidència documental.	Implementar una acció de generació d'evidència associada a aquella tasca, i si escau fer arribar aquesta evidència a l'agent que genera la sèrie primària.
	La documentació de la tasca s'està arxivant en una altra sèrie documental.	Opció 1: generar un segon exemplar de l'evidència que genera la tasca i fer-lo arribar a l'agent que genera la sèrie primària.
		Opció 2: substituir l'evidència que genera la tasca per una còpia, i fer arribar l'original a l'agent que genera la sèrie primària.
Una evidència essencial de la sèrie primària no té caràcter d'original.	Un agent de la mateixa organització està retenint l'exemplar original, i està enviant la còpia a l'agent generador de la sèrie primària.	Indicar a l'agent que reté l'original que inverteixi la seva acció: es quedi la còpia, i faci arribar l'original a l'agent que genera la sèrie primària.
	Un agent d'una altra organització està retenint l'exemplar original, i està enviant la còpia a l'agent generador de la sèrie primària.	Sol·licitar a l'agent de l'altra organització que generi un segon exemplar original de l'evidència i el faci arribar a l'agent que genera la sèrie primària.
	La tasca s'està documentant sense cap procediment d'autenticació.	Implementar un procediment d'autenticació en la generació de l'evidència dins l'execució de la tasca.

3.3.4 Pas 3: deshibridació

Un procés dadificat és sempre un procés digital. El procés pot contemplar puntualment la recepció de documents en suports no digitals, però l'emmagatzematge final dels documents sempre serà en format digital.

En aquesta fase de la transformació, atès que encara no s'introdueix cap sistema d'informació, sinó que només es fan modificacions sobre la naturalesa i característiques de les evidències ja existents, no es plantejarà la supressió de cap evidència, ni tampoc es plantejarà cap actuació més enllà del nucli informacional del procés: les evidències essencials de les sèries primàries.

3.3.4.1 Identificar les evidències en formats no digitals i transformar-ne el format

El primer pas serà, doncs, centrar l'atenció en les evidències essencials de les sèries primàries que no estiguin en format digital.

Actualment l'únic format no digital que es pot trobar en els processos administratius és el format paper. L'ús de microformes, suports analògics, etc. es pot considerar ja pràcticament descartat. Així doncs, totes aquelles evidències essencials de les sèries primàries que estiguin en format paper es transformaran a format de document-objecte digital, d'acord amb el següent patró de transformació:

Figura 3-18: Pautes d'actuació per transformar el format de les evidències a digital

Format de l'evidència original	Procés de transformació	Format de l'evidència transformada
Evidència creada externament (sense control possible sobre el procés de creació) i rebuda originalment en paper.	Implantar un sistema de digitalització en origen, per tal de transformar l'evidència a un document digitalitzat.	Format de fitxer resultant d'un procés de digitalització (PDF, JPEG, TIFF, etc.).
Evidència creada externament en format digital, impresa i rebuda originalment en paper.	Habilitar un canal per a la recepció telemàtica de l'evidència en format digital.	En funció del format en què es generi. En aquest cas es pot mantenir com a document-objecte, o pot ser ja una evidència basada en dades.
Evidència creada internament en format digital, impresa i rebuda en paper per l'agent productor de la sèrie primària.	Indicar a l'agent que genera l'evidència que faci arribar a l'agent productor de la sèrie primària l'original digital.	En funció del format en què es generi. Habitualment es mantindrà com a document-objecte.

3.3.5 Pas 4: automatització

Un cop es disposa d'una col·lecció d'evidències íntegrament digitals, es pot començar a treballar en la seva automatització.

Aquesta automatització pot prendre una primera forma, que no modifica substancialment la forma de les evidències. Consisteix en mantenir totes les evidències com a objectes diferenciats, però reduint les agrupacions documentals únicament a aquelles que són essencials per al procés (sèries primàries). Per tant, el tret destacat d'aquest pas és que se suprimeixen les sèries secundàries, però no la capacitat d'actuació dels agents que les generen. Per aconseguir-ho, se substitueix la possessió o custòdia directa de les evidències per part d'aquests agents, per un accés a les evidències de les sèries primàries, passant de múltiples expedients a l'expedient compartit. I per a que això sigui possible, cal implementar una tecnologia que permeti gestionar de forma àgil un accés concorrent a documents compartits.

Aquesta tecnologia és la que popularment es coneix com a gestor documental, o *Electronic Document Management System* (EDMS), la qual té múltiples manifestacions en el mercat, que van des dels entorns de col·laboració amb funcionalitats de compartició de documents, a elements d'infraestructura que estan pensats per actuar com una capa per sota de les aplicacions de gestió. En aquesta guia utilitzarem genèricament la denominació *gestor documental* per referir-nos a aquesta tipologia d'eines.

3.3.5.1 Incorporar un gestor documental com a destí de totes les evidències essencials

El primer que cal fer és preveure la utilització d'un gestor documental per hostatjar totes les evidències essencials de les sèries primàries del procés. I, en conseqüència, se substituirà l'indret d'emmagatzematge de tots els documents d'aquestes sèries pel gestor documental.

Un cop fet això, es valoraran les dues casuístiques amb impacte procedimental que s'expliquen en els dos apartats següents.

3.3.5.2 Valorar el tractament de la missatgeria

Totes les evidències de tramesa d'informació entre agents (bàsicament correus electrònics) hauran de ser objecte de valoració. La decisió que caldrà prendre és si

aquestes evidències s'han de convertir en evidències-objecte que resideixin al gestor documental, o bé han de ser només avisos conforme els agents ja tenen accés a alguna evidència del sistema.

En el primer cas, el canvi procedimental a fer consistirà en buscar la forma per a que la generació de l'evidència doni com a resultat un document-objecte formalitzat. En el segon cas, s'haurà de substituir l'evidència per missatgeria del sistema. Aquesta missatgeria serà volàtil, en el sentit que no generarà cap objecte a conservar, però no irrellevant, atès que haurà de generar una traça de comunicació associada a l'evidència comunicada en forma de metadades o pista d'auditoria.

El criteri principal per a la discriminació d'ambdues opcions és el següent:

- Les evidències de tramesa amb contingut propi (text significatiu vinculat al desenvolupament del procés), especialment quan no s'acompanyin d'un altre document, hauran de ser tractades com a evidències-objecte.
- Les evidències de tramesa amb un contingut irrellevant, o que facin esment només al document al qual acompanyen, hauran de ser tractades com a missatgeria del sistema.

En el primer cas, la modificació que es farà al procediment consistirà en substituir una acció de tramesa d'un correu electrònic (o altre tipus de missatge) per una acció de dipòsit d'una evidència-objecte en l'expedient de la sèrie primària. Aquesta acció, prèvia creació de l'evidència, la farà l'agent emissor del missatge. Si escau, aquest agent també haurà d'activar la generació del corresponent avís a l'agent receptor.

En el segon cas, la modificació consistirà en substituir l'acció de tramesa del correu electrònic (o altre tipus de missatge) per una acció d'anotació de comentari o generació d'avís en l'expedient electrònic de la sèrie primària, acció associada a una evidència o directament a l'expedient en funció del cas. La imminència de l'avís al receptor dependrà de les funcionalitats que permeti el gestor documental utilitzat.

3.3.5.3 Suprimir les redundàncies documentals en les evidències essencials

La segona casuística a tractar és l'existència de còpies o exemplars duplicats d'una evidència dins el circuit documental del procés. En qualsevol tramesa d'informació, sempre es generen múltiples exemplars de la unitat d'informació que es transmet: l'exemplar que arxiva l'agent emissor, com a testimoni de la tramesa, i els exemplars que reben cadascun dels agents receptors. Això és així perquè la necessitat de documentar l'activitat durant el procés es genera a nivell d'agent, no d'evidència. És cada agent qui té la necessitat d'explicar la seva activitat dins el procés, i és també cada

agent qui té capacitat d'actuar amb accions documentals. El procés en conjunt és una entelèquia, no fa accions. La clau en aquest cas és poder explicar l'activitat de cada agent dins el procés sense haver de multiplicar les manifestacions o exemplars de les evidències mitjançant còpies.

La manera d'eliminar aquestes còpies, que formaran part habitualment de les sèries documentals secundàries derivades de la sèrie primària, és fer possible que tots els agents puguin accedir a l'evidència original que ha generat les còpies o exemplars, i en puguin disposar quan n'hagin de menester. La substitució de la possessió (d'un exemplar de l'evidència) per l'accés (a l'exemplar únic original de l'evidència) és la forma d'eliminar duplicats d'evidències essencials d'un procés. I això només és possible quan les evidències resideixen en una plataforma que permet l'accés concurrent als documents, i que n'acota els privilegis d'ús. Sense plataforma o programari de gestió documental, doncs, és pràcticament impossible eliminar les còpies o exemplars múltiples del circuit documental d'un procés.

3.3.5.4 Modificar el procés per a que generi un circuit d'expedient únic

Fent un repàs de la casuística abans esmentada, podem identificar el següent patró de transformacions a realitzar sobre el circuit evidencial del procés:

Taula 3-6: Pautes d'actuació per a generar un circuit d'expedient únic

Evidència en situació inicial	Transformació
Evidència en forma de document-objecte generada per un agent, i tramesa a l'agent productor de la sèrie primària. Genera dos exemplars, el de l'agent emissor i el de l'agent receptor.	Es modificarà el procediment per tal que l'agent emissor creï l'evidència-objecte directament dins la sèrie primària. Per aquest motiu es donaran permisos d'escriptura a l'agent emissor sobre l'expedient compartit de la sèrie primària, que podran estar subjectes a temporalitat. Eventualment es podran activar avisos del sistema per notificar a l'agent receptor la creació i dipòsit de l'evidència.
Evidència en forma de document-objecte tramesa per l'agent productor de la sèrie primària a un altre agent de l'organització. Genera dos exemplars, el de l'agent emissor i el de l'agent receptor.	Es donaran permisos de lectura a l'agent receptor sobre l'expedient compartit de la sèrie primària, que finalitzaran com a molt aviat al tancament de l'expedient. En dipositar l'agent productor l'evidència a l'expedient, dispararà un avís del sistema a l'agent receptor per notificar-li la disponibilitat d'un nou document.

Evidència en situació inicial	Transformació
Evidència no rellevant (missatge) generada per un agent, i tramesa a l'agent productor de la sèrie primària. Genera dos exemplars, habitualment de correu electrònic: el de l'agent emissor i el de l'agent receptor.	Es donaran permisos d'anotació o generació d'avísos a l'agent emissor sobre l'expedient compartit de la sèrie primària. Alternativament es mantindrà la possibilitat de trametre correus electrònics informatius, que no seran objecte de tractament ni d'arxiu.
Evidència no rellevant (missatge) tramesa per l'agent productor de la sèrie primària a un altre agent de l'organització. Genera dos exemplars, habitualment de correu electrònic: el de l'agent emissor i el de l'agent receptor.	En aquest cas no caldrà donar permisos a l'agent receptor sobre l'expedient compartit de la sèrie primària, només sobre la plataforma en conjunt. L'agent emissor utilitzarà la missatgeria del sistema per generar avisos i comunicacions adreçades a l'agent receptor. Alternativament es mantindrà la possibilitat de trametre correus electrònics informatius, que no seran objecte de tractament ni d'arxiu.

3.3.5.5 Implementar el procediment en un gestor documental

La metodologia descrita s'ocupa únicament de la definició del circuit documental del procés, i de la caracterització del conjunt d'evidències que el componen. Arribats a aquest punt s'ha reduït un circuit que contenia:

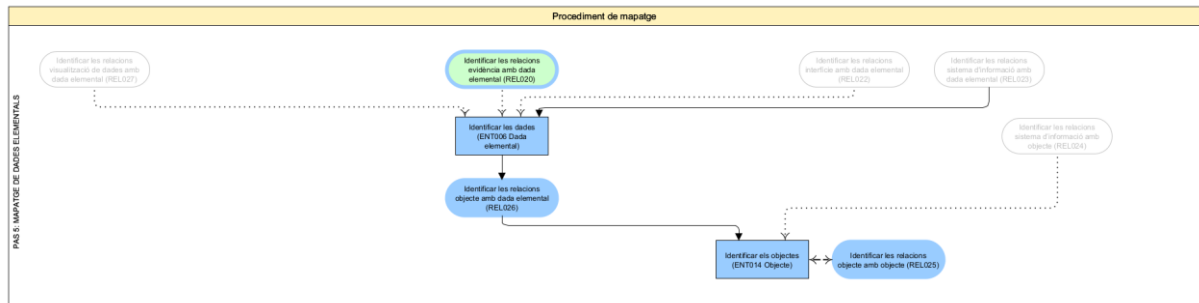
Taula 3-7: Resultat documental de la transformació a un circuit d'expedient únic

Situació inicial	Situació un cop transformat a docucèntric
N sèries documentals secundàries	0 sèries documentals secundàries
N sèries documentals primàries, amb les següents característiques: <ul style="list-style-type: none"> • Amb part de les evidències essencials del procés • Amb part de les evidències amb valor d'original • Amb evidències essencials i d'enllaç • Amb evidències en diferents formats • Amb evidències en diferents ubicacions o àmbits de custòdia 	N sèries documentals primàries, amb les següents característiques: <ul style="list-style-type: none"> • Amb totes les evidències essencials del procés • Amb totes les evidències amb valor d'original • Només amb evidències essencials • Amb totes les evidències en format digital • Amb totes les evidències sobre l'àmbit de custòdia dels productors de les sèries primàries

Però aquí no finalitza l'automatització. Sobre la base d'aquest circuit s'inicia un projecte tecnològic de desenvolupament i implantació d'una solució de gestió documental, que implicarà la definició d'una estructura d'objectes (tipus d'expedient i agrupacions), model de dades (metadades dels objectes i agrupacions), assignació de permisos mitjançant llistes de control d'accés i capacitats (ACL), i diversos tipus de flux de treball, que poden incloure avisos i notificacions, flux de tasques o regles de cicle de vida. Aquesta definició queda fora de l'abast d'aquesta guia, no només pel fet que es tracta d'un procediment de desenvolupament tecnològic sobradament normalitzat, sinó perquè aquesta guia vol posar l'èmfasi en el punt següent: la dadificació del circuit evidencial.

3.3.6 Pas 5: dadificació

Figura 3-19: Estructura per a la dadificació del procés



3.3.6.1 Incorporar un sistema d'informació com a destí de totes les evidències essencials

Arribats al punt anterior, hem aconseguit un substrat evidencial compost per 1 o N sèries documentals primàries, les quals contenen totes les evidències essencials que documenten el procés, en format digital i amb atributs d'evidència original, i sota l'àmbit lògic de custòdia del productor de cada sèrie. Les tasques o accions del procés que no generen evidències essencials han estat substituïdes per missatgeria del sistema, preferentment volàtil (no arxivable). Aquest escenari d'òptim documental és indispensable per poder iniciar qualsevol procés d'automatització realment efectiu.

En el punt anterior s'ha exposat com es porta a terme aquesta automatització quan es basa en evidències-objecte. Anem a veure a continuació com podem orientar-ne el disseny a dades. Per tant, substituïrem en el circuit evidencial el destí d'arxiu dels documents essencials, que fins ara era un gestor documental, per un sistema d'informació encara indeterminat.

La interacció amb els usuaris en una aplicació informàtica de gestió administrativa es porta a terme a través de pantalles que visualitzen dades en forma de formularis. Cada pantalla o visualització de dades té un objectiu concret, proporciona a l'usuari unes dades de treball, i en pot recollir unes altres, generades pel propi usuari o per l'activitat de l'usuari dins el sistema (captura automàtica). Per tant, es tracta de visualitzacions funcionals que agreguen dades ja disponibles en el sistema amb variables d'informació que s'empenaran amb dades facilitades per l'activitat de l'usuari.

La principal diferència amb un circuit documental basat en documents-objecte (docucèntric) és aquesta: en un circuit docucèntric, cada tasca pot tenir evidències d'entrada (documents que es reben), i evidències de sortida (documents que es generen). No és habitual que sobre els documents rebuts es facin modificacions, atès

que habitualment seran documents ja definitius i amb caràcter d'original. Però en un circuit basat en dades (datacèntric), les evidències d'entrada i de sortida d'una tasca es poden trobar físicament fusionades en una mateixa pantalla o formulari, on una part de les dades ja venen informades (prèvia cerca o de forma directa), i una altra part s'informen (es creen) com a resultat d'aquella tasca. Com tractarem conceptualment aquest fet en el disseny del circuit datacèntric?

Si ens limitem a transposar cada evidència-objecte en una pantalla de dades, estarem malbaratant bona part dels guanys procedimentals que pot aportar un circuit basat en dades. Hem d'entendre cada evidència com una interacció, en la qual el que realment ens interessa és saber quines dades es visualitzen i quines es generen. Per tant, abans de decidir quines evidències del tipus pantalla de dades inclourà el procés dadificat, ens centrarem en dissenyar el model de dades que haurà de donar suport a aquest procés.¹⁸¹

3.3.6.2 Construir el model (lògic) de dades del sistema d'informació que donarà suport al procés

El nucli de qualsevol sistema d'informació és el conjunt de dades que utilitzarà. Per poder-les utilitzar d'una determinada manera, aquestes dades han d'estar estructurades seguint unes regles determinades. Aquestes regles són, doncs, les que converteixen (conceptualment) un conjunt de dades en una base de dades, i en permeten una automatització eficient mitjançant sistemes gestors de bases de dades. Habitualment, en els sistemes de negoci o dissenyats de bell nou, estructurarem les dades d'acord amb una lògica relacional, però en funció del tipus de dades a gestionar, del volum de dades, de la velocitat amb la qual s'hagin de processar i de la precisió que calgui aplicar en la seva recuperació podrem utilitzar altres lògiques (desnormalització en sistemes BI, jerarquització, orientació a objectes, etc.).

La tecnologia informàtica ha desenvolupat nombroses metodologies per a l'estructuració de conjunts de dades i per a la seva modelització, amb l'objectiu de dissenyar una base de dades. Aquestes metodologies es basen en recollir informació per entendre el context d'utilització d'un conjunt de dades i dissenyar-ne una estructura el més eficient possible. La metodologia que exposem aquí se centra únicament en el moment inicial del disseny d'una base de dades: la identificació del conjunt de dades que hauran de ser objecte de gestió i posterior estructuració. Ens mourem doncs sempre en el pla lògic, sense entrar

¹⁸¹ Aquesta metodologia de dadificació es pot utilitzar únicament amb finalitats descriptives, atès que no sempre es donen les condicions de context per poder portar a terme una dadificació completa. En aquest cas, només s'executaran els passos que es descriuen en els apartats 3.3.6.2.1, 3.3.6.2.2 i 3.3.6.2.3.

en els elements propis del disseny físic de la base de dades (diccionaris de dades, taules de relació, etc.).

La construcció d'aquest model lògic es realitza en dues etapes. La **primera**, centrada únicament en el procés objecte d'anàlisi, consta de les següents operacions tècniques:

1. Anàlisi individual de la informació de cada evidència, on haurem de detectar, identificar i acotar sistemàticament totes les ocurrències o aparicions d'una dada, el que en terminologia APS s'anomena variables d'informació,¹⁸² i categoritzar-les convenientment.
2. Construcció del mapa seqüencial de la informació del procés, base per a la seva posterior automatització. Per aconseguir-ho, elaborarem una matriu (el que APS anomena graella de variables) amb la representació de les variables de cada evidència, seguint la seqüència del circuit documental del procés (ordre d'aparició de les evidències en el procés).

La **segona**, un cop es disposa de la matriu detallada del procés, té per objectiu mapar aquesta relació amb la informació preexistent a l'organització, que podem trobar dispersa en evidències d'altres processos, però també consolidada en sistemes d'informació ja existents, i per tant en una forma estructurada i reutilitzable.¹⁸³ Amb aquesta operació s'obre la porta a optimitzar l'ús de la informació per part del procés, i a integrar-lo en el mapa de dades global de l'organització. Les operacions que es realitzen són les següents:

1. Mapatge de la matriu de variables del procés amb la matriu de variables global de l'organització, amb una atenció especial a les dades consolidades en altres processos ja automatitzats que puguin ser utilitzades com a dades mestres.
2. Valoració de les variables ja existents a la matriu de variables global que poden substituir la recollida de variables d'entrada del procés, substituint aquestes variables per variables directament utilitzables (més endavant les anomenarem "importades").
3. Valoració de possibles noves variables que, incorporades al procés, puguin proporcionar als agents una millor informació per a la presa de decisions, i

¹⁸² Una variable d'informació no és una dada, sinó el contenidor d'una dada. La mateixa dada pot estar continguda en diferents variables. La variable és part del document, mentre que la dada és part del conjunt d'informació de l'organització. En el nostre model ens hi referim amb el terme dada elemental (per definir la dada) i ocurrència de la dada (la relació d'una dada amb una evidència o un sistema d'informació). El concepte de variable el considerem equivalent al que nosaltres proposem d'ocurrència, i l'utilitzem en aquesta guia per respecte a la denominació original utilitzada en la metodologia APS.

¹⁸³ També en podríem dir interoperable. Incloent, si estan reconegudes com a tals, les dades mestres de l'organització.

incorporació d'aquestes variables a la matriu de variables i a les interaccions del procés on puguin aportar més.

4. Substitució de les evidències documentals anteriors per noves evidències d'interacció amb sistemes d'informació, agrupades en funció de cada activitat del procés.

Amb aquestes operacions haurem aconseguit un model de dades lògic del procés, integrat en el model global de l'organització.¹⁸⁴ Utilitzant aquesta nova matriu com a model de dades, podrem passar al desenvolupament de la solució informàtica que doni suport al nou procés. Aquest modelat físic ja es durà a terme atenent a la tecnologia concreta en què s'hagi d'implementar aquesta solució.

3.3.6.2.1 Identificació dels objectes descrits

Aquesta operació consisteix en identificar les entitats que apareixen en el procés i relacionar-les formant una sentència. Les entitats obtingudes seran de dos tipus:

- **Entitats transaccionals**, habitualment només una per procés. Correspon a l'acció que es realitza durant el procés. El propi procés, o bé la seva/es sèrie/s documental/s primària/es pot donar nom a aquesta entitat, que és sempre una acció, mai un objecte o cosa, i per tant està sempre subjecte a temporalitat (inici i fi). En el nostre model les representem amb el valor "acció" informat en l'atribut **ENT014-002 Tipus d'objecte**.
- **Entitats de referència**. Corresponen habitualment als agents que participen en el procés, i especialment als agents externs que són destinataris de l'acció realitzada. Si l'acció realitzada ha generat algun tipus de producte, aquest també serà una entitat de referència. En tots els casos seran coses del món real, no accions. En el nostre model les representem amb el valor "cosa" informat en l'atribut **ENT014-002 Tipus d'objecte**.

Per a la identificació dels objectes, es proposa tractar el procés com una oració gramatical. El plantejament és similar al que s'utilitza per a la definició de triplets semàntics en el llenguatge RDF, o per al modelat dimensional en bases de dades desnormalitzades.¹⁸⁵ El primer que caldrà fer és identificar la declaració del document resolutori del procés, que serà el verb de l'oració (predicat en terminologia RDF). A continuació s'identificaran els diferents agents o ens que conformin tant el subjecte com l'objecte de l'oració (els objectes sobre els que recau l'acció). El verb generarà

¹⁸⁴ A nivell físic pot ser que haguem detectat dades mestres que no es consoliden en cap sistema d'informació, implantacions incompletes, etc. L'objectiu és obtenir una integració de les dades a nivell lògic, i utilitzar-la després per anar optimitzant el mapa de sistemes d'informació de l'organització.

¹⁸⁵ Pels referents vegeu l'apartat 2.2.3.3 del capítol 2, i també l'apartat 3.1.2.4 d'aquest capítol.

habitualment un objecte transaccional, del tipus “acció”, mentre que el/s subjecte/s i els objectes/s seran habitualment objectes del tipus “cosa”.

Per exemple, en el cas dels programes d'innovació educativa, un centre educatiu, a través del seu director/a, sol·licita a l'Administració poder participar d'un programa d'innovació. L'Administració resol i adjudica o denega la participació en el programa. Les entitats d'aquest procés seran:

- Entitats transaccionals:
 - La participació en un programa d'innovació educativa.
- Entitats de referència:
 - El centre educatiu.
 - El treballador/a del centre educatiu amb rol de director/a.
 - L'Administració educativa.
 - El programa d'innovació educativa.

Un cop construïda la sentència, es buscarà a l'inventari d'objectes si aquests ja han estat identificats prèviament, i en cas contrari es donaran d'alta els nous objectes. Aquestes entitats es donaran d'alta com instàncies de l'entitat **ENT014 Objecte**, informant els següents atributs:

- **ENT014-001 Nom**
- **ENT014-002 Tipus**

També s'informaran, a partir d'aquesta detecció, les relacions entre els diferents objectes, a través de la relació **REL025 objecte amb objecte**.

Un cop identificades les entitats del procés, es procedirà al buidatge de les variables de cada evidència. Les evidències s'ordenaran per l'ordre d'aparició en el procediment, i s'aniran buidant les seves variables seguint els dos passos que s'expliquen a continuació.

3.3.6.2.2 Extracció de les variables de cada evidència

Aquesta operació consistirà en fer el buidatge de totes les variables d'informació de cada evidència del circuit documental, incloent en aquest buidatge absolutament totes les evidències (les essencials i les d'enllaç o missatgeria). Per a fer això, obtindrem exemplars de cadascuna de les evidències del procés, completament emplenades i en format definitiu. Aquests exemplars poden ser documents-objecte, pantalles de visualització de dades, missatges i avisos, o qualsevol altre format d'evidència.

Les evidències poden incloure ja algun nivell d'estructuració de la informació. Això ens permetrà distingir entre tres grups d'evidències:

- a) **Evidències totalment estructurades**, on cada variable ja està enunciada amb un títol o literal descriptiu. És el cas dels formularis i de les pantalles de visualització de dades.
- b) **Evidències parcialment estructurades**, on les variables estan majoritàriament predeterminades, però no sempre apareixen ni ho fan en la mateixa disposició. En aquest cas trobarem variables amb un major contingut informatiu, com poden ser descripcions, observacions, signatures, etc. És el cas dels missatges o avisos, i dels documents pautats pel llenguatge administratiu o jurídic (informes, resolucions, etc.).
- c) **Evidències no estructurades**, on la informació pot aparèixer en qualsevol disposició no previsible. És el cas de narracions, declaracions, imatges fixes o en moviment, gravacions de veu, etc. Aquestes evidències són el principal repte per a l'obtenció de variables.

Els processos no es desenvolupen sempre seguint el mateix flux de tramitació. En uns casos poden seguir un camí (per exemple una beca que es concedeix) i en altres casos un camí diferent (per exemple una beca que es denega). També es poden interrompre en diversos punts (per exemple un estudiant que desisteix d'una sol·licitud de beca). Això té com a resultat que els expedients d'un mateix procés no sempre contindran les mateixes evidències: n'hi haurà de més constants en la seva aparició, i d'altres que només apareixeran ocasionalment.

Per mitigar l'impacte d'aquesta situació, els exemplars a buidar s'obtindran de diversos expedients del procés, que hagin finalitzat de formes diferents i que s'hagin tramitat en moments diferents. Com més àmplia sigui la mostra utilitzada, més complet serà el mapa de variables que s'obtindrà.

Per al buidatge, s'ordenaran les evidències per aquest ordre: primer les més estructurades, i finalment les menys estructurades o de contingut menys previsible. I es farà així pel següent motiu. En tot procés, les evidències comparteixen sovint entre elles un nombre elevat de variables d'informació; les mateixes variables apareixen sovint en diversos documents. Atès que la identificació de variables és més fàcil i precisa com més estructurada està l'evidència (ja que la variable sempre està descrita amb més detall), buidar primer els documents més estructurats ens donarà un coneixement del context que ens ajudarà a identificar les variables (sovint les mateixes) en aquells documents on estan menys descrites i són més difícils d'identificar (els no estructurats).

No podem evitar trobar-nos amb documents incomplets, formularis mal dissenyats o informacions contradictòries. Per aquest motiu, quan detectem possibles anomalies ho validarem obtenint més exemplars de l'evidència sospitosa. Intentarem doncs sempre

obtenir evidències d'expedients i moments diferents per poder detectar millor possibles regularitats en l'aparició de variables.

Per a l'acceptació o rebuig de les variables dubtoses, utilitzarem la pauta fixada per la metodologia APS: donarem prioritat a aquelles variables que apareguin en el darrer terç del temps cobert amb tots els exemplars de cada evidència, i que es detectin com a mínim en el 50% dels exemplars analitzats. Aquelles variables que sempre apareguin en blanc o no informades en tots els exemplars de les evidències del procés valorarem excloure-les d'aquesta extracció.

Si els dubtes són importants, parlarem també amb els diferents agents del procés, per tal de poder detectar possibles variables, determinants per a certes decisions que es prenguin en el curs del procés, que només es transmetin per via oral i no quedin recollides en cap evidència del procés.

Tota evidència té una forma externa, visible, que conté la informació que ha de comunicar, i pot contenir també informació implícita, invisible en la seva forma externa però accessible a través de l'anàlisi d'aquell fitxer (sovint amb la condició de metadades implícites). Per tant, el buidatge de les evidències inclourà:

1. El buidatge de la capçalera, on habitualment trobarem variables com l'agent productor (logotip), codis de classificació, dates de registre, etc.
2. El buidatge del cos, on habitualment trobarem les variables més relacionades amb el propi procés, aquelles que podem qualificar de variables transaccionals (com per exemple un estat del tràmit, una data de lliurament d'un producte, etc.). Una de les variables més importants d'aquesta part es deriva del propi tipus de document, el que s'anomena "declaració" del document.
3. El buidatge del peu, on habitualment trobarem les variables d'autenticació (signatures o validacions d'agents, segellats, etc.), datacions (tòpiques i cronològiques), explicacions o compromisos, etc.
4. El buidatge de les variables implícites, on habitualment trobarem informació d'agents, dates rellevants, versions, etc.

La identificació de les variables no sempre és senzilla. Sovint les variables estan emmascarades en formes no convencionals de representació. Per aquest motiu, és important prestar atenció a alguns detalls significatius:

- A. En els documents en paper o amb forma externa convencional, les datacions, tòpiques i cronològiques, es poden trobar en forma explícita, però també integrades dins de segells o marques diverses, o també en identificacions de versions. En els documents electrònics les dates es poden trobar també formant part de signatures electròniques basades en certificat (tinguin o no la consideració de segell de temps), o bé en metadades de l'objecte.

- B. Els agents poden trobar-se també en forma explícita, o bé representats en capçaleres i logotips, signatures, etc. En funció de la ubicació del logotip dins el document, l'agent podrà ser considerat productor, autor, destinatari, etc.
- C. Un dels elements més importants a identificar és la *declaració* del document. S'entén per *declaració* aquell verb que condensa i descriu l'acció que es vol portar a terme amb aquell document o evidència. La declaració determina el tipus documental de forma directa. Per tant, una evidència que contingui la paraula "Sol·licito" serà considerada una sol·licitud, mentre que una altra evidència, tot i que tingui exactament els mateixos elements formals i variables d'informació, que contingui la paraula "Resolc" serà considerada una resolució. La declaració no sempre és una variable explícita, i pot trobar-se:
 - a. Com un element destacat de la part declarativa de l'evidència.
 - b. Com a part del títol de l'evidència (en el cas de formularis o documents estructurats).
 - c. O bé derivar-se del redactat del cos del document.

Una mateixa variable pot aparèixer en diferents evidències amb noms, formes i semàntiques diferents. Per tal de construir un model normalitzat que reflecteixi correctament la matriu de variables del procés, un cop fet el buidatge haurem d'harmonitzar les variables per saber que no estem recollint com a dues variables diferents el que realment correspon a la mateixa dada. Per evitar duplicitats i imprecisions es realitzaran doncs dues harmonitzacions consecutives:

1. En primer lloc, es construirà la relació de variables de totes les manifestacions o exemplars analitzats d'una mateixa evidència, aplicant per a la seva tria l'algoritme anteriorment descrit (donarem prioritat a aquelles variables que apareguin en el darrer terç del temps cobert amb tots els exemplars de cada evidència, i que es detectin com a mínim en el 50% dels exemplars analitzats).
2. I en segon lloc, un cop obtinguda la relació de variables de cada evidència del procés, es construirà una segona llista, única per al procés, amb la fusió de totes les variables obtingudes de cadascuna de les evidències del procés. Amb això s'obindrà la relació de variables d'aquell procés, la qual la metodologia APS anomena "**Diccionari de variables del procés**".
3. Per completar l'harmonització es combinaran ordenacions del diccionari de variables per nom de la variable, per tipus i per objecte, per tal d'eliminar duplicitats i depurar la categorització (informació descriptiva) de cada variable o dada.

Aquesta harmonització a una llista única ens permetrà consultar el mapa global de variables de l'organització, recollit a l'entitat **ENT006 Dada elemental**, i veure si aquella dada ja existeix. En cas que ja existeixi, procedirem a assignar aquella dada elemental

a totes les evidències del procés on apareix, a través de la relació **REL020 evidència amb dada elemental**. En cas que la dada elemental s'hagi detectat per primer cop, aleshores es crearà una nova instància de l'entitat **ENT006 Dada elemental** per aquella dada, i s'informaran els atributs que la caracteritzen:¹⁸⁶

- **ENT006-001 Identificador**
- **ENT006-002 Nom**
- **ENT006-003 Tipus**
- **ENT006-004 Descripció**
- **ENT006-005 Regles de validació**
- **ENT006-006 Nivell de seguretat RGPD**
- **ENT006-007 Disponible per interoperabilitat externa**

Tant en un cas com en l'altre, la dada elemental ve determinada per l'objecte al qual descriu. La identificació prèvia dels objectes del procés que s'ha fet en el punt anterior servirà per assignar cada dada a l'objecte adient, a través de la relació **REL026 objecte amb dada elemental**.

Aquesta assignació no és trivial, i una mala assignació pot afectar negativament el modelat posterior. Es tracta de raonar realment què descriu cada variable, no en el seu context sinó de forma directa. Per exemple, en un procés de compra d'un producte, on cada instància del procés seria una "Compra", la variable "Nom i cognoms" que apareix a l'evidència "Sol·licitud de compra", i que probablement s'arrossega en altres evidències del circuit, no pertany directament a l'objecte "Compra", sinó que pertany a l'objecte "Client". És a dir, és el *Nom i cognoms – Del client – D'una compra*. Per tant, l'ens descrit per la variable "Nom i cognoms" no serà "Compra" sinó "Client". En canvi, la variable "Data de compra" només la podem atribuir a l'objecte "Compra", no hi haurà dubte en aquest cas. Així, doncs, la primera variable serà una dada que descriu una cosa del món real, i per tant candidata a dada reutilitzable (dada mestra), cosa la qual implica que podrà aparèixer com a dada importada al procés. En canvi, la segona variable serà una dada autòctona, que es crea de nou amb l'execució del procés, i que descriu directament l'acció en què consisteix el procés: la compra.¹⁸⁷

¹⁸⁶ Per un detall de com informar aquests atributs, vegeu la Guia d'inventari i mapatge (**GMAP**).

¹⁸⁷ En els apartats següents s'amplien els conceptes de variable autòctona i importada, i els tipus d'objectes o entitats de dades.

3.3.6.2.3 *Construcció de la matriu de variables del procés*

Com s'ha comentat anteriorment, les evidències d'un procés no són ens independents. Les evidències adquireixen la seva significació completa (i el seu valor d'actuació o vigència) quan es relacionen amb el seu context, un context que és doblement funcional i documental:

- Funcional pel fet que tota evidència és la materialització informacional d'una acció que s'ha esdevingut o d'una tasca que s'ha desenvolupat. I aquestes accions o tasques es relacionen de forma natural les unes amb les altres, en una relació de causa-efecte que, en el seu conjunt, conforma el que s'anomena procés.
- I documental pel fet que cada acció o tasca genera una o més evidències, que hereten les relacions funcionals que mantenen entre elles les accions o tasques que les han generades. I aquestes relacions funcionals queden recollides de forma material en les variables d'informació que conté cada evidència, i especialment en la repetició en diferents evidències de les mateixes dades o variables d'informació.

Per tant, tota evidència compartirà amb altres evidències del mateix procés un conjunt de dades o variables d'informació, i aquesta compartició determinarà els vincles entre les evidències.

Per representar aquesta relació, el que es farà és disposar totes les variables del diccionari de variables del procés en una matriu de doble entrada, on l'eix d'abscisses representarà les evidències del procés, i l'eix d'ordenades les variables d'informació. Aquesta visualització permetrà analitzar de forma global el conjunt d'informació que genera el procés. I el principal element d'anàlisi serà un nou atribut amb el qual es caracteritzarà la relació de cada dada o variable amb cada evidència del procés: l'origen del valor de la variable.

D'acord amb això, podrem identificar dos tipus de relació entre la variable o dada i l'evidència:

- Parlarem de **variables autòctones** quan la variable d'informació no existeixi en una forma informada en el moment en què s'inicia la creació de l'evidència, i s'empleni amb un valor com a resultat de l'acció o tasca que genera aquella evidència. Aquestes variables, que la metodologia APS anomena també "variables d'entrada", es poden emplenar amb una acció manual per part d'un usuari o bé de forma automàtica a partir d'algun tipus de càlcul, fórmula o operació realitzada dins la pròpia evidència.¹⁸⁸

¹⁸⁸ Aquest darrer conjunt, les variables emplenades de forma automàtica, la metodologia APS les anomena "variables d'operació". En aquest model les considerarem també variables d'entrada o autòctones.

- Parlarem de **variables importades** quan la variable d'informació ja contingui un valor quan s'inicia la creació de l'evidència.¹⁸⁹ El terme importada descriu el fet que el valor de la variable no es genera durant la creació de l'evidència, sinó que s'ha generat en una acció o tasca anterior, i pot aparèixer preempletat per facilitar informació de suport al creador de l'evidència. Les variables importades d'una evidència són variables que ja han aparegut en evidències anteriors del procés o, en el cas de la primera evidència del procés, que s'han creat en evidències de processos anteriors.

Només les evidències basades en dades, és a dir les que tenen al darrera un sistema d'informació, poden visualitzar variables preempenades. Per tant, en marcar l'origen del valor de la variable aplicarem per defecte aquesta regla:¹⁹⁰

- A. Les evidències de tipus "document" (evidència-objecte) tindran sempre totes les variables marcades com autòctones, llevat que hagin estat generades de forma automatitzada a partir d'un sistema d'informació (mitjançant l'ús de plantilles).
- B. Les evidències de tipus "dada" (interfícies d'entrada i sortida de dades) es marcaran com autòctones o importades, en funció de si adquireixen o modifiquen un valor en aquesta evidència (autòctones) o bé simplement visualitzen un valor informat anteriorment (importades).

Per poder categoritzar adequadament aquestes relacions, és indispensable que les evidències en la matriu de variables segueixin en aquest punt una ordenació estrictament procedimental, d'acord amb la seva aparició en el procés. Aquesta categorització es materialitzarà a través de l'atribut **REL020-001 Tipus d'ocurrència de la relació REL020 evidència amb dada elemental**.

¹⁸⁹ La metodologia APS les anomena "variables de ròssec", però amb un canvi substancial en el concepte que s'explica en la nota següent. Hem mantingut l'ús del terme ròssec com a sinònim de reutilització en els indicadors SER-16, SER-17, PRO-17 i PRO-18 proposats al capítol 5.

¹⁹⁰ Aquest és un canvi conceptual en relació amb la metodologia APS, la qual considera només com a variable autòctona la primera aparició d'una variable dins el procés. APS prioritza la perspectiva de procés, de manera que tota variable té una primera aparició com autòctona i la resta com importada ("de ròssec"). En aquesta metodologia s'ha modificat aquest criteri, marcant com autòctona la variable sempre que obliga l'usuari a informar novament un valor per aquella variable, i com importada quan simplement opera com variable de lectura. D'aquesta manera es pot utilitzar aquesta categorització com un indicador de nivell d'integració del procés. L'òptim d'un procés automatitzat és que cada variable només s'informi un sol cop, i per tant cada nova actualització d'una variable ja existent en el procés s'ha de considerar una disfunció o inoperància del procés.

3.3.6.2.4 Mapatge amb la matriu de variables global de l'organització

El concepte de matriu de variables, que hem vist fins ara associat a un procés en concret, és extrapolable al conjunt de processos d'una organització. De la mateixa manera que podem disposar en una matriu totes les variables que apareixen als documents d'un procés, també podem elaborar una matriu amb les variables de tots els processos d'una organització. Això no és una solució nova, ja que fa temps que les empreses i institucions construeixen els seus sistemes d'informació al voltant d'un nucli comú de dades. Aquestes dades comunes a tots els sistemes, i que per tant tenen un major potencial de reutilització, s'anomenen "dades mestres", i la seva definició obeeix a l'anomenat principi de la dada única (Butler, 2011). Segons aquest principi, cada dada ha d'estar emmagatzemada només una vegada a l'organització, i per assegurar-ne la integritat qualsevol ús de la dada (que pot implicar-ne la replicació) ha d'assegurar la sincronització amb la dada original.¹⁹¹

Als efectes de modelat de dades, però, qualsevol dada és reutilitzable dins el conjunt de processos de l'organització. El potencial de reutilització interna és doncs el que determina el seu tractament com a dades mestres. En aquest sentit, podem diferenciar dos tipus de dades:

1. Aquelles dades que descriuen coses del món real que existeixen amb independència de la relació amb l'organització, com poden ser persones, centres educatius, països, etc. Són les que tenen un potencial major de reutilització, i per tant les que habitualment se segregen en forma de dades mestres.
2. Aquelles dades que descriuen coses que són resultat directe dels processos de l'organització. Es tracta habitualment d'accions substantivades, com poden ser una auditoria, una subvenció o una compra. Habitualment s'anomenen dades transaccionals, i no acostumen a ser considerades objecte d'un sistema de gestió de dades mestres.

¹⁹¹ Les solucions tecnològiques que permeten implementar el concepte de dada única són múltiples (Kusuma et al., 2019). Atès que la connexió punt a punt entre tots els sistemes de l'organització és inviable o extremadament costosa, una solució és habilitar un *bus* o *hub* al qual es connecten els sistemes per intercanviar dades de forma sincronitzada, solució que sovint es complementa amb un orquestrador dels processos d'intercanvi (*Enterprise Application Integration*). Però el mètode més habitual, si es disposa dels recursos suficients, és la creació d'un repositori centralitzat al qual es connecten els sistemes per obtenir i/o dipositar aquelles dades que més es reutilitzen. Aquest repositori pot contenir des de les dades més comunes en qualsevol procés de l'organització (geogràfiques, temporals, etc.), passant per les dades de referència bàsiques de l'organització (personal, clients, etc.), fins a les dades generades pels diferents processos o tràmits (dades transaccionals consolidades). La gestió d'aquests repositoris es fa habitualment amb solucions anomenades *Master Data Management* (MDM). Per a més informació, vegeu l'apartat 2.2.2.1.2 del capítol 2.

Ambdós tipus de dades, a mesura que s'acumulen, van construir el nucli informacional de l'organització, aquell conjunt d'informació única i original que descriu el funcionament i l'activitat de l'organització.

Per depurar la matriu de variables d'un procés, el primer que farem serà comprovar quines variables de la matriu del procés ja existeixen al mapa de variables del conjunt de l'organització. Per facilitar aquesta tasca recuperarem el diccionari de variables del procés, que hem elaborat en l'apartat 3.3.6.2.2.

Abans de donar el diccionari per tancat, farem una darrera operació. Partirem de la idea que el resultat de cada execució d'un procés, i de qualsevol de les sèries documentals que se'n deriven, és un expedient. En aquest cas un expedient format únicament per dades, les que recull el diccionari de variables. Però en la seva forma objectual, tot expedient es defineix per unes metadades, habitualment convencionals i sovint formalitzades amb estàndards o normes de descripció (data d'inici, data de fi, número d'expedient, estat, etc.). Aquestes metadades a nivell d'expedient poden ser explícites, i haver-les detectat a través de les variables d'alguna de les evidències analitzades, o bé ser implícites. En aquest segon cas el buidatge de variables de les evidències no les haurà detectades. Per resoldre aquesta situació, es verificarà que el diccionari de variables del procés inclou totes les metadades d'un expedient tipus, i en cas contrari s'afegiran al diccionari per assegurar la seva correspondència amb l'estàndard que s'apliqui en aquella organització.

A continuació, un cop completat, confrontarem el diccionari de variables del procés amb el mapa global de variables de l'organització, i marcarem aquelles variables que ja existeixen reconegudes com a dades mestres, tant si es tracta d'informació de referència com transaccional. El determinant per aquesta designació de dades mestres serà la naturalesa de l'objecte descrit. I sobre la base dels objectes que descriuen les dades del diccionari de variables del procés, crearem el model de dades del procés.

3.3.6.2.5 Modificació de les variables d'entrada del procés

Un cop fixades les variables de referència del procés, analitzarem quines d'aquestes variables apareixen a la matriu de variables del procés com a variables autòctones. Anirem doncs a cercar, dins del procés, la primera referència on apareix cada variable, i en el cas que es tracti d'una variable autòctona, la marcarem per a la seva posterior transformació. L'objectiu és transformar en variables importades totes aquelles variables autòctones del procés que ja existeixen al mapa de dades de l'organització (que ja tenen un valor per a aquell procés), i que no requereixen ser objecte d'actualització en aquell procés. Això ens permetrà agilitzar el procés i descarregar a l'agent productor de cada evidència de la tasca d'informar dades de què l'organització ja disposa, reduint alhora la possibilitat de cometre errors informant dades que ja han estat consolidades prèviament.

Un cop identificades, transformarem cadascuna d'aquestes variables a nivell de disseny en una variable de sortida, és a dir, en variables importades a partir dels sistemes que les contenen, amb els quals s'hauran d'establir les connexions corresponents. Això implicarà revisar el circuit del procés per veure com aquelles dades de què ja es disposa a l'organització poden incorporar-se a una evidència de nova creació, incloent els corresponents processos de cerca, recuperació i validació.

Aquesta substitució de variables d'entrada per variables de sortida obligarà a connectar nombroses evidències del procés amb sistemes d'informació ja existents i no vinculats prèviament amb aquest procés, la qual cosa en molts casos pot implicar modificacions substancials en el format i estructura d'aquestes evidències. Això pot comportar la creació de noves evidències, la fusió d'evidències ja existents, i la substitució de moltes evidències per noves evidències d'interacció amb base de dades, agrupades en funció de cada activitat del procés (pantalles d'entrada de dades o formularis electrònics, pantalles o entorns de visualització de dades, etc.).

3.3.6.2.6 Afegint valor al procés

Però la gran aportació que aquesta operació pot fer al procés no és únicament simplificar i reduir la quantitat de variables que els usuaris hauran d'aportar al procés a través de les seves evidències, sinó sobretot ampliar el nombre de variables a què els usuaris tindran accés al procés moment de prendre decisions. En aquest cas, el punt d'atenció serà novament l'objecte descrit per cada variable.

Així doncs, si una variable descriu un objecte del qual l'organització té més informació, podem fer que l'usuari, en el moment d'informar una nova dada, pugui tenir accés al nombre més gran de variables sobre aquest objecte, visualitzant aquestes dades a la pròpia interfície d'entrada. Recuperant l'exemple anterior, imaginem que hem convertit el nom i els cognoms del client en una variable importada, en lloc de ser una variable d'entrada, i que el recuperem a partir d'un número identificador aportat per l'usuari (el DNI/NIF, per exemple). Enlloc de limitar-nos a mostrar el nom i els cognoms del client, podem afegir a l'evidència, per exemple, l'adreça postal i el número de telèfon (obtinguts d'una font de referència, el fitxer de clients), que ajudaran l'usuari a poder contactar amb el client. Però no només això. També podem afegir informació de fonts transaccionals, com ara les dates i contingut de les últimes compres (del propi fitxer de compres), o fins i tot dades d'altres fonts transaccionals no directament relacionades amb el procés de compres, com ara l'historial de reclamacions del client o els seus impagaments. D'aquesta manera estarem donant una informació de context a l'usuari molt superior a la que podria obtenir del mateix procés de compres, i estarem contribuint a augmentar de manera molt notable la precisió i l'encert en les decisions que prengui per a la continuació del procés en una direcció o una altra. Si aquesta informació és realment útil, extrem que verificarem parlant amb l'usuari, la incorporarem a la matriu de variables

del procés, i en conseqüència a l'evidència inicial del procés, en forma de variables importades.¹⁹²

El fet d'haver categoritzat les variables per objecte descrit serà clau per poder realitzar aquesta operació. Per tant ordenarem les variables per objecte descrit, i per a cada variable buscarem en el mapa global de variables quines altres variables descriuen també el mateix objecte, i poden ser d'utilitat per a les decisions que s'han de prendre en el procés objecte d'automatització.

3.3.6.3 Modificar el procés per a que generi el circuit documental dadificat

Un cop disposem del model de dades del futur sistema d'informació, podrem començar a dissenyar els seus components funcionals. Aquests són bàsicament les pantalles de visualització i edició de dades, i les operacions de tractament i flux que s'hauran de realitzar en *back-office*. En aquest moment podem recuperar el circuit evidencial optimitzat en el pas anterior, i modificar-lo d'acord amb la pauta següent:

Taula 3-8: Pautes per a la transformació del circuit documental a datacèntric

Situació en el circuit docucèntric	Modificació en el circuit datacèntric
Evidències originals generades pels agents externs.	Formularis d'entrada de dades accessibles externament.
Evidències originals generades com a part del procés.	Pantalles de visualització de dades (vistes), amb permisos d'edició (alta, modificació, baixa) per als usuaris que estiguin executant cada tasca.
Missatges de tramesa d'evidències a agents interns diferents del productor.	Pantalles de visualització de dades (vistes), amb permisos de lectura per als usuaris que hagin de rebre el document (pot incloure un flux o missatge d'avís).
Missatges de tramesa d'evidències a agents externs (amb accés al sistema).	Pantalles de visualització de dades (vistes), amb permisos de lectura temporals i d'abast limitat per als usuaris que hagin de rebre el document (pot incloure un flux o missatge d'avís).

¹⁹² Aquesta possibilitat és la que obre la porta a l'ús sistemàtic de visualitzacions interactives de dades com a part de la gestió, proposta que es desenvolupa en l'objectiu 6 de les conclusions (apartat 8.3.6 del capítol 8).

Situació en el circuit docucèntric	Modificació en el circuit datacèntric
Evidències originals rebudes dels agents externs i no generades per ells.	Evidències-objecte autenticades, rebudes en annex als formularis d'entrada.
Missatges de tramesa d'evidències a agents externs (sense accés al sistema).	Evidències-objecte autenticades, generades a demanda i trameses a l'agent extern.

La pervivència en el circuit d'evidències-objecte farà necessari, a fi de poder gestionar i arxivar aquestes evidències, incorporar al sistema d'informació un component o aplicació de gestió documental, o bé funcionalitats que n'emulin el funcionament.

3.3.6.4 Implementar el procediment en un sistema d'informació

A partir d'aquí, ja disposem de tots els elements necessaris per començar a desenvolupar el nou sistema d'informació. El modelat físic, les formes de les evidències i el procés de desenvolupament del sistema dependran de la metodologia i tecnologia que es decideixi utilitzar per a la seva construcció. En aquest cas, els factors que incidiran dependran de l'estratègia tecnològica de l'organització, i de qüestions d'oportunitat i recursos disponibles.

Com hem comentat a l'inici d'aquest apartat, el pas del model lògic a un model físic de dades, i la seva posterior implementació en base de dades, com també el desenvolupament d'una solució tecnològica, són elements que queden fora de l'abast d'aquesta metodologia. No obstant, la derivació a partir d'un procés basat en documents a un model lògic de dades representatiu del procés permet fer aportacions interessants per a la definició de l'arquitectura de sistemes d'informació d'una organització, assenyalant espais d'informació on es fa necessària la consolidació tecnològica de conjunts de dades mestres, fent evident la necessitat d'interconnexió de determinats sistemes, o bé optimitzant la inversió en desenvolupaments tecnològics, en funció de la rellevància, volum i impacte de cada procés sobre el mapa de dades de l'organització.

3.4 Annex: proposta de sistematització de la recollida d'informació

El procés d'elaboració dels mapes, i en especial del mapa de processos, requereix disposar d'informació el més detallada possible sobre l'estructura, les activitats i els actius d'informació de l'organització. Per obtenir aquesta informació, i aconseguir que sigui exhaustiva i de qualitat, cal accedir a fonts d'informació internes i externes i obtenir dades mitjançant diferents tipus d'actuacions.¹⁹³ Però no només això: també cal fer-ho d'una forma prou sistemàtica per a que el procés es pugui portar a terme en un temps i amb uns costos ajustats als beneficis del projecte. Per aquest motiu, com a complement operatiu de la metodologia exposada, es descriuen a continuació una sèrie de pautes per ordenar i sistematitzar aquesta recollida d'informació.¹⁹⁴

3.4.1 Anàlisi prèvia

Abans d'iniciar el procés, és convenient fer una anàlisi de les característiques de l'organització i del context en què opera. Les variables d'informació que cal considerar són les següents:

- Característiques constitutives, especialment si es tracta d'un organisme públic o privat, el règim jurídic, els ens de control i responsabilitat, etc.
- Recursos humans, tecnològics i econòmics, dimensionament de l'organització, sistemes d'obtenció de recursos econòmics, distribució territorial, etc.
- Usuaris o clients actuals i potencials, demandants de serveis, obligacions i restriccions relacionades amb el tipus d'usuaris, etc.
- Necessitats i objectius a curt i llarg termini, existència de plans estratègics, integració en plans de caràcter general, etc.
- Funcions, activitats i procediments, existència de reglamentacions, manuals de procediment, etc.

¹⁹³ És important que les persones que el realitzen, siguin de la pròpia organització o externes (consultors), tinguin habilitats per a l'anàlisi i la modelització de realitats complexes, per a la comunicació oral i escrita, i per al maneig de fonts d'informació de tot tipus, i que disposin des del principi del procés del suport directe i explícit de la direcció de l'organització per fer aquesta tasca. Sense aquest suport, serà molt difícil mantenir la intensitat de treball necessària per completar adequadament el projecte.

¹⁹⁴ Les quals es consideren bàsiques per poder avaluar els objectius instrumentals que es plantegen en l'apartat 1.3.2 del capítol 1.

- Sistemes existents i actuacions prèvies, incloent-hi sistemes de gestió documental, sistemes de gestió de qualitat, sistemes informàtics, i també iniciatives anteriors de millora en l'àmbit de procediments, documentació, qualitat i tecnologia.

Aquest estudi previ permet, no sols detectar les fonts d'informació inicials, sinó també planificar la cerca de fonts internes més específiques, i també de fonts externes necessàries per a la comprensió correcta de les activitats de l'organització i del seu context. Així, per exemple, per a qualsevol empresa que gestioni dades de clients, cal que l'analista conegui amb cert detall els requisits que a aquestes empreses els imposa la normativa de protecció de dades de caràcter personal, i encara que aquesta normativa no estigui descrita en cap document propi de l'organització, és convenient incloure-la també al registre de fonts d'informació que cal construir.

Aquest estudi previ també permet determinar l'abast de la tasca d'anàlisi que s'emprendrà, especialment quan el projecte se circumscriu només a una part o unitat d'una organització més gran. En aquest cas, proporciona la visió global necessària per centrar l'anàlisi en la part de l'organització objecte del projecte sense obviar la pertinença d'aquesta unitat a un conjunt més gran i l'impacte d'aquesta dependència en el context funcional.

3.4.2 Inventari i priorització de fonts d'informació

El primer pas consisteix a identificar i recopilar les fonts d'informació que s'utilitzaran per elaborar el sistema de classificació. Una sèrie de pautes poden ajudar a realitzar el treball de camp per obtenir aquestes fonts.

Aquest pas consisteix a registrar de la manera més exhaustiva possible cadascuna de les fonts d'informació que es localitzin, amb la finalitat no només de controlar-ne la disponibilitat i el buidatge, sinó també de disposar d'una referència única per a la citació durant l'elaboració del sistema de classificació. Aquest registre és una peça clau per a l'aprofitament posterior de la informació obtinguda, ja que és molt important que, per a cada funció o activitat, s'indiqui la font o fonts de les que s'ha obtingut, per tal de disposar de criteris per a la seva validació o depuració durant l'elaboració del sistema de classificació.

Aquest registre ha d'incloure informació sobre:

- Les fonts documentals, internes o externes, formals o informals, la disponibilitat física o en línia i el nivell de buidatge.
- Les fonts personals: unitats administratives i persones de les quals se n'ha obtingut informació.

- Les activitats de recollida d'informació. Es consideren fonts les notes de reunions, les entrevistes, els qüestionaris i les notes d'observacions directes.

Aquesta sistematització és necessària per assegurar l'exhaustivitat de la recollida d'informació, potenciar el rigor en la seva anàlisi, detectar duplicitats o sinonímies, i evitar la confusió que es pot generar durant les iteracions i les revisions fins a la finalització dels diferents mapes.

Quins tipus de fonts d'informació cal obtenir? Es pot establir la següent tipologia:

1. Fonts personals

1.1. Fonts directes

1.1.1. Entrevistes formals

1.1.2. Qüestionaris

1.2. Fonts indirectes

1.2.1. Entrevistes informals

1.2.2. Observació directa

2. Fonts documentals

2.1. Fonts directes

2.1.1. Documents públics (publicats)

2.1.2. Documents sobre constitució i planificació

2.1.3. Documents de recapitulació i memòria

2.1.4. Normativa i procediment

2.1.5. Projectes de millora

2.2. Fonts indirectes:

2.2.1. Formularis, models de documents i altres pautes formals

2.2.2. Sistemes d'informació i models de dades

2.2.3. Documentació en general

Analitzarem amb més detall aquestes fonts, i especialment el tipus d'informació que podem obtenir de cadascuna.

3.4.2.1 Fonts personals

Les fonts personals són el primer mètode d'obtenció d'informació sobre les funcions i les activitats de l'organització analitzada. Proporcionen informació de primera mà, actualitzada i fidel a la realitat, però que requereix un procés de contrastació i filtre per eliminar les aportacions subjectives. Es divideixen en:

3.4.2.1.1 Entrevistes formals

Una de les formes bàsiques d'obtenir informació sobre les funcions i activitats és mitjançant reunions o entrevistes.

El procés d'entrevistes es desenvolupa de la manera següent:

3.4.2.1.1.1 Determinar/decidir el tipus d'entrevista i a qui entrevistar

Hi ha diferents mètodes d'entrevista. En funció de l'estructura de l'organització i dels objectius de cada entrevista, l'analista decidirà el mètode òptim en cada cas:

- L'entrevista **individual** proporciona informació de detall sobre funcions i activitats, i permet aprofundir-ne el desenvolupament.
- L'entrevista **en grup** proporciona informació general i també transversal i permet contrastar dades o relacionar informacions que afecten més d'una persona o unitat orgànica.

Siguin individuals o en grup, és important que les entrevistes incloguin tots els nivells de personal de l'organització (directiu, tècnic i operatiu). Cada nivell proporciona informació valuosa:

- Les entrevistes amb el **nivell directiu** no ajudaran a conèixer els objectius i el perquè de les activitats de l'organització. Són especialment útils a les primeres fases de l'anàlisi, i és preferible que siguin individuals.
- Les entrevistes amb el **nivell tècnic** ens aportaran informació sobre com s'estructuren les funcions de l'organització en activitats, quins són els requisits de context en el desenvolupament d'aquestes activitats, i perquè es descomponen en determinats procediments. Aquestes entrevistes es poden fer inicialment en grup, i completar-les de forma individual. En cas que l'organització disposi d'un arxiver, la seva aportació és imprescindible per conèixer el disseny i la problemàtica dels sistemes de gestió documental.

- I finalment les entrevistes amb el **nivell operatiu** ens permetran conèixer (i contrastar) la manera com es duen a terme les activitats, els problemes d'índole pràctica, les formes documentals concretes que es generen (per exemple models de documents), i fins a quin punt les activitats es duen a terme d'acord amb els procediments definits.

L'organigrama de l'organització és una eina valuosa per identificar àrees i interlocutors, i així planificar les entrevistes. És important entrevistar com a mínim un representant de cada àrea operativa de l'organització, preferiblement el responsable de l'àrea o un tècnic amb experiència. Aquesta persona pot assenyalar la necessitat d'ampliar l'espectre de persones a entrevistar, i proporcionar nous contactes i fonts d'informació.

3.4.2.1.1.2 Programar les entrevistes

Per minimitzar l'afectació sobre el personal entrevistat, és recomanable comunicar la durada les entrevistes, el seu nombre, i dins de quin període es faran. És important que aquest període no coincideixi amb èpoques de sobrecàrrega de feina per a les persones entrevistades. Per tant, cal elaborar un calendari detallat d'entrevistes, adaptat a les necessitats dels entrevistats, i consensuar-lo amb ells, considerant sempre un cert marge de variabilitat per incloure entrevistes ajornades o ampliacions d'entrevistes a causa de nova informació recollida.

És recomanable que les entrevistes no excedeixin l'hora de durada. És útil per a l'analista reservar-se un temps després de cada entrevista per compilar les notes, i no demorar molt més aquesta tasca, ja que és el moment en què més recent és la informació i més dades es poden completar.

3.4.2.1.1.3 Preparar les preguntes

El propòsit de les entrevistes és obtenir informació de detall sobre les característiques i les activitats de l'organització, i validar la informació obtinguda d'altres fonts, especialment de les fonts documentals. Per aquest motiu es pressuposa que, abans d'iniciar el procés d'entrevistes, l'analista ha adquirit un coneixement bàsic de les funcions i l'estructura de l'organització i que ha iniciat el buidatge de les fonts documentals.

Per preparar les primeres entrevistes, doncs, hi ha documents bàsics que cal consultar, ja que proporcionen una perspectiva general de l'organització i permeten optimitzar la recollida d'informació personal. Aquests són:

- La memòria anual.
- L'organigrama.

- La pàgina Web, la Intranet i els comptes en xarxes socials.
- La legislació i normativa tècnica més general.

Amb independència del tipus d'entrevista, i del possible ús complementari de qüestionaris, per assegurar que l'entrevista serà productiva i orientar l'entrevistat sobre la informació que ens interessa, cal preparar una llista de preguntes que poden incloure informació sobre:

- Les àrees de responsabilitat (activitats) de l'entrevistat
- Com es relacionen aquestes àrees amb altres i amb el conjunt d'activitats de l'organització
- Les tasques concretes que es realitzen per dur a terme aquestes activitats
- El possible ús de terminologia específica
- Afectacions legals i normatives de les activitats
- Documentació generada pel desenvolupament de les activitats i eventuales duplicitats o còpies, incloent-hi els sistemes d'informació
- Relacions documentals amb altres àrees de responsabilitat
- Valoració sobre la durada útil de la documentació generada
- Validació i ampliació de les fonts d'informació recopilades/utilitzades

3.4.2.1.1.4 Conduir les entrevistes

És important fer les entrevistes de forma relaxada, en un lloc adient i respectant els horaris especificats. En iniciar l'entrevista s'ha d'informar l'entrevistat dels objectius de l'entrevista, i en acabar explicar-vos les properes passes i eventuales noves entrevistes.

3.4.2.1.1.5 Recollir els resultats de les entrevistes

Com ja hem comentat, és recomanable que les notes de l'entrevista es compilin de forma immediata a la finalització, quan la informació encara està fresca a la memòria. També cal prendre les notes de forma estructurada, seguint el guió de l'entrevista. Si l'entrevistat no té cap inconvenient, també pot ser útil gravar les entrevistes.

Totes les entrevistes es registraran al registre de fonts d'informació i rebran un codi identificador. Durant la resta de l'anàlisi es tractaran amb caràcter general com a fonts d'informació.

Per obtenir la informació en una seqüència lògica i poder ordenar còmodament la informació obtinguda es recomana seguir la pauta següent:

- **Entrevista inicial amb el nivell directiu**, amb l'objectiu d'identificar els objectius estratègics i les funcions principals, conèixer l'estructura i les àrees de l'organització, i obtenir la designació d'interlocutors per a cadascuna de les àrees.
- **Entrevista en grup amb els nivells directiu, tècnic i operatiu** de cadascuna de les àrees de l'organització, per conèixer les activitats principals i sol·licitar fonts documentals.
- **Entrevistes més detallades amb cada nivell de cada àrea de l'organització.** Aquestes entrevistes poden arribar a ser personals en cas que es detectin desacords evidents o inexactituds entre diferents fonts d'informació. Aquestes entrevistes han de tenir caràcter iteratiu, per tal de contrastar la informació proporcionada mitjançant tècniques assertives, i s'han de prolongar durant tot el procés d'elaboració del sistema de classificació.

3.4.2.1.2 Qüestionaris

Les entrevistes es poden acompanyar de qüestionaris, preferentment oberts o semioberts. Quan es disposi d'un temps limitat per part del personal de les àrees, els qüestionaris s'utilitzaran per preparar i optimitzar les entrevistes. Els qüestionaris també es poden utilitzar durant les entrevistes, a manera de guió, cosa que alhora facilita l'estructuració posterior de la informació recollida. No es recomana l'ús exclusiu de qüestionaris en substitució de les entrevistes, per la dificultat de contrastar la informació obtinguda.

3.4.2.1.3 Entrevistes informals

Sovint la informació que s'obté durant les entrevistes formals, especialment si es fan en grup, pot estar condicionada per múltiples factors (jerarquia, desavinences internes, desconfiança o desconeixement del projecte). Sempre que sigui possible, es recomana aprofitar qualsevol oportunitat informal (és a dir no programada) per obtenir més informació o contrastar aquella de què ja es disposa. És important que, al conjunt del projecte, aquestes fonts informals rebin el mateix tractament i registre que la resta de fonts d'informació.

3.4.2.1.4 Observació directa

Cal no rebutjar cap oportunitat d'obtenir informació. Determinada informació potser no serà mai proporcionada per cap membre de l'organització, especialment aquella que es considera massa òbvia (per exemple, el canvi de direcció dels associats d'una associació). Un analista experimentat pot obtenir aquesta informació mitjançant l'observació directa de les instal·lacions de l'organització, o bé de l'activitat en temps real dels seus membres. Com en el cas anterior, és important que aquestes observacions rebin el mateix tractament i registre que la resta de fonts d'informació.

3.4.2.2 Fonts documentals

L'anàlisi de documentació és el segon mètode per obtenir informació funcional sobre l'organització. Cada tipus de document proporciona informació sobre un aspecte determinat de les funcions de l'organització, i necessita un tractament determinat.

Abans de procedir al buidatge, es recomana que totes les fonts d'informació obtingudes s'ordenin seguint l'ordre que s'indica en aquest apartat. Aquest ordre intenta garantir que, en primer lloc, s'obté la informació de caràcter més general (objectius i funcions), i que progressivament es va detallant aquesta informació fins a obtenir-se la més específica (activitats i procediments). Aquesta seqüència ajuda l'analista a disposar sempre del context suficient per comprendre de manera òptima la informació que va analitzant i orientar-se en el procés de sintetitzar i sistematitzar el substrat funcional de l'organització analitzada.

Aquestes fonts es divideixen en:

3.4.2.2.1 Documents públics

Proporcionen una visió general de l'organització i les seves activitats, i aporten informació sobre la imatge que l'organització vol projectar a l'exterior, així com sobre els usuaris/clients o els destinataris dels serveis i productes de l'organització. Això permet identificar els objectius i les funcions més importants per a l'organització, i també les activitats amb destinataris externs, que probablement es correspondran amb els processos clau. Documents d'aquest tipus són els butlletins digitals, publicacions impreses, publicacions en xarxes socials, i especialment la pàgina Web.

Aquestes fonts van dirigides a una audiència general i no pressuposen un coneixement detallat de l'organització, per això són fonts ideals per a una aproximació inicial a les funcions i activitats de l'organització. A més, permeten conèixer l'existència de fonts més específiques i remetre's-hi. En aquest sentit, és especialment important prestar atenció

a la pàgina web, tant per la seva informació general sobre l'estructura i les activitats de l'organització, com sobre els seus serveis clau, especialment quan s'ofereixen en línia a través de la pàgina web.

3.4.2.2 Documents sobre constitució i planificació

Proporcionen informació sobre els objectius estratègics i les funcions d'alt nivell de l'organització. Són documents d'aquest tipus els estatuts i les escriptures fundacionals, els plans estratègics i altres declaracions de voluntat, i també les descripcions de l'estructura orgànica.

Els documents constitutius habitualment venen fixats legalment. Els organismes públics es poden crear mitjançant disposicions normatives publicades a diaris oficials. En el cas dels organismes privats amb finalitat de lucre, es creen jurídicament d'acord amb el que estableix el Dret mercantil. Els organismes privats sense finalitat de lucre es constitueixen mitjançant l'aprovació d'uns estatuts i la seva inscripció en el registre corresponent.

Mereixen una atenció especial els documents de planificació, ja que ubiquen les funcions de l'organització dins una estratègia global, i dins aquest context els donen sentit. Com que són documents dirigits bàsicament a un públic intern, proporcionen informació essencial que permet vincular cada àrea amb una esfera de responsabilitat o activitat, i vincular així mateix aquestes activitats entre elles.

3.4.2.3 Documents de recapitulació i memòria

Proporcionen informació sobre com s'ha donat resposta als objectius de l'organització mitjançant activitats concretes. Inclouen les memòries anuals, els resums diaris d'activitat i les estadístiques.

Les memòries anuals són una font d'informació especialment important, ja que pretenen oferir una visió global i alhora detallada de l'estructura, els objectius i les activitats de l'organització. Aquesta informació sovint està quantificada, i a través d'aquests indicadors es pot conèixer també el pes específic de cada activitat sobre el conjunt de l'organització.

3.4.2.4 Normativa i procediment

Proporcionen informació sobre el marc d'activitat de l'organització, i com afecta aquest marc el desenvolupament de les activitats. Permet identificar totes les activitats

reglades, i inclou la normativa legislativa i reglamentària pública, la normativa tècnica (estàndards), les polítiques i els manuals de procediment.

La legislació es considera una font externa a l'organització. A més de la normativa legal constitutiva o emanada de l'organització (en cas que es tracti d'un organisme públic), cal considerar en tots els casos la legislació de caràcter general aplicable a totes les organitzacions de característiques i àmbit d'activitat similars.

L'analista no s'ha d'interessar únicament per la normativa vigent. Bona part de les activitats de l'organització s'expliquen per normes que ja han estat modificades. Val la pena, per tant, establir la necessària relació de precedència entre les normes esmentades.

Els estàndards són una altra font important d'informació. Es pot tractar de codis de bones pràctiques, de protocols, d'estàndards tècnics o industrials, i poden ser voluntàriament adoptats per l'organització o de compliment obligat.

Les polítiques contenen els motius pels quals les organitzacions s'organitzen de manera determinada i com realitzen les seves activitats. Els procediments, sovint compilats en manuals, desenvolupen la manera de dur a terme les activitats, indicant-ne les fases i les accions. Ambdós documents poden estar relacionats amb una funció, una part d'una funció, aspectes de diverses funcions, o amb el conjunt d'activitats de l'organització. En tots els casos són l'eina bàsica per conèixer amb detall les activitats, ubicar-les en el context funcional i descompondre-les en activitats menors o transaccions.

3.4.2.2.5 Projectes de millora

Poden proporcionar informació ja elaborada sobre les activitats de l'organització (mapes de processos), i també sobre actuacions anteriors que condicionin la informació que podem obtenir de les fonts presencials. És important estudiar les condicions de desenvolupament dels projectes esmentats per valorar adequadament la vigència i el valor de la informació que proporcionen, i la seva possible reutilització a nivell d'anàlisi funcional.

3.4.2.2.6 Formularis, models de documents i altres pautes formals

Proporcionen informació sobre els procediments i les transaccions concrets, juntament amb les variables d'informació gestionades per cada procediment (existència de documents recapitulatius, sèries relacionades, etc.). També tenen un potencial important per detectar activitats i procediments operatius no detectats mitjançant les fonts formals.

3.4.2.2.7 Sistemes d'informació i models de dades

Els sistemes d'informació i aplicacions de gestió cada vegada són més una font d'informació valuósíssima per detectar activitats i procediments, i també per analitzar variables d'informació. Cal tenir present el model de dades de cada aplicació, els formularis d'entrada, els informes i les estadístiques generats, els intercanvis de dades entre aplicacions i el patró d'utilització per part dels usuaris.

3.4.2.2.8 Documentació en general

Finalment, és convenient recórrer a la consulta de documents i expedients d'exemple de cadascuna de les activitats i procediments, analitzar-ne la composició, l'ús de formularis i models de documents, les fases, etc. En aquest sentit, és extremadament important tenir molt en compte l'agrupació natural en expedients realitzada per les pròpies àrees gestores, els seus criteris de nominalització, codificacions sectorials que s'hagi emprat, i la freqüència d'aparició de determinades agrupacions i anomalies.

El fet que l'organització disposi de sistemes anteriors de gestió documental es reflectirà en aquesta organització natural dels documents i proporcionarà informació molt valuosa sobre funcions i activitats. A això hi contribuirà l'existència de sistemes de classificació anteriors.

Capítol 4. Eina de recollida i model de dades

4.1 Introducció

Per a la recollida de la informació s'ha desenvolupat una eina específica. Aquesta eina s'ha plantejat com un catàleg global de tots els actius funcionals i tecnològics que participen de la gestió de la informació en una organització. Inclou les funcionalitats bàsiques d'alta, modificació i baixa, juntament amb funcionalitats d'importació, exportació i connectivitat amb plataformes externes. A continuació es descriuen les característiques funcionals d'aquesta eina.

Atès que el primer bloc d'informació que es va incorporar a l'eina fou el mapa de processos, en un primer moment aquesta eina va rebre la denominació de Catàleg de Processos, i els usuaris s'hi han continuat referint amb l'acrònim CATPROC. Malgrat que el nom no reflecteix el contingut i abast complet de l'eina, l'hem mantingut en el marc d'aquesta tesi perquè a dia d'avui és el nom amb el qual és coneguda internament al Consorci d'Educació de Barcelona.

4.2 Versions

La primera versió de l'eina, encara sense denominació, es basava en un full de càlcul implementat en *Microsoft Excel* en el qual es podien informar els processos, els agents, les sèries documentals, els sistemes d'informació, les evidències i les dades que apareixien en cada evidència. Aquesta versió es va utilitzar des de 2016 per a la recollida inicial d'informació.

La primera versió estable es va desenvolupar en modalitat SaaS, en un *tenant* de llicència pròpia de *Microsoft Sharepoint*, però amb les dades emmagatzemades en *Microsoft SQL Server*. Aquesta versió, denominada **CATPROC 1.0**, va permetre consolidar la informació recollida inicialment i fer-ne a partir d'aquell moment manteniment. En un primer moment proporcionava funcionalitats d'alta, baixa i modificació de cadascuna de les entitats, juntament amb el manteniment dels diccionaris associats als atributs amb validació. L'eina incorporava les funcionalitats de cerca i traçabilitat pròpies de *Microsoft Sharepoint*.

Posteriorment, i amb el suport d'un desenvolupador,¹⁹⁵ es va construir una capa de serveis que va proporcionar funcionalitats de cerca avançada, i especialment va incloure les funcionalitats d'exportació (en els diferents formats que es detallen en els apartats següents) i d'importació i processament de fluxgrames.

Aquesta versió inicial es va enriquir posteriorment amb un evolutiu que va afectar el model de dades, amb la incorporació de les entitats "Quadre de comandament" (*dashboard*) i "Interfície de connexió" entre sistemes d'informació (atributs de la relació sistema a sistema). Aquest evolutiu es va incorporar l'any 2020, quan l'eina va passar a ser un recurs de consulta habitual de l'equip de Sistemes i Processos del Consorci d'Educació de Barcelona.

La segona versió (**CATPROC 2.0**) fou requeriment de l'Àrea de Tecnologies de la Informació i les Comunicacions del Consorci d'Educació de Barcelona l'any 2022, en esdevenir aquesta eina un recurs cada cop més utilitzat per al govern de la informació al Consorci. La segona versió va implicar un canvi tant a nivell de la tecnologia de base com del servidor on estava hostatjada. L'eina es va reprogramar en tota la seva capa de serveis, i es va migrar a un servidor *Microsoft Azure*, abandonant en aquest cas *Microsoft Sharepoint* com a interfície d'usuari però mantenint la base de dades en el seu entorn original (*Microsoft SQL Server*). Aquest és actualment l'entorn on està consolidada i es manté la informació recollida en el curs del projecte.

La voluntat del projecte és mantenir en paral·lel una versió de l'eina íntegrament desenvolupada sobre *Microsoft Sharepoint* (**CATPROC 3.0**), basada en l'ús de llistes de *Sharepoint* i amb funcionalitats limitades al manteniment de la informació, que pugui ser utilitzada per qualsevol organització que tingui interès en aplicar el model desenvolupat.

L'entorn d'explotació i anàlisi, que ofereix la consulta dels indicadors generats en el curs del projecte a través de diferents consoles o *dashboards* interactius, s'ha desenvolupat íntegrament en *Microsoft Power BI*, i es connecta directament a *Microsoft SQL Server* (per a la versió 2.0) i a *Microsoft Sharepoint* (per a la versió 3.0).

A continuació es detallen les característiques funcionals de la versió 2.0, atès que és la que conté el major nombre de funcionalitats implementades.

¹⁹⁵ Es va comptar amb el suport del Sr. Carles Drudis, de l'empresa NEC2, qui va desenvolupar, d'acord amb les indicacions de l'autor, la capa de serveis i els successius evolutius de l'eina.

4.3 Funcionalitats

4.3.1 Funcionalitats d'administració

Inicialment CATPROC 1.0 fou utilitzat únicament per l'autor, i amb la finalitat de recollir dades per al projecte. Les càrregues inicials de dades es van fer directament contra base de dades, a partir de fitxers *Microsoft Excel*, i es va completar la recollida a través de la interfície d'usuari. Però a partir de CATPROC 2.0 l'eina ha començat a ser utilitzada per diversos equips del Consorci d'Educació de Barcelona per a la consulta i manteniment de la informació. Per aquest motiu, CATPROC 2.0 incorpora una gestió d'usuaris bàsica, que inclou l'alta i la baixa d'usuaris. L'accés a l'eina requereix autenticació a través d'un usuari facilitat per l'administrador, qui també disposa de l'opció de donar d'alta i de baixa nous usuaris.¹⁹⁶

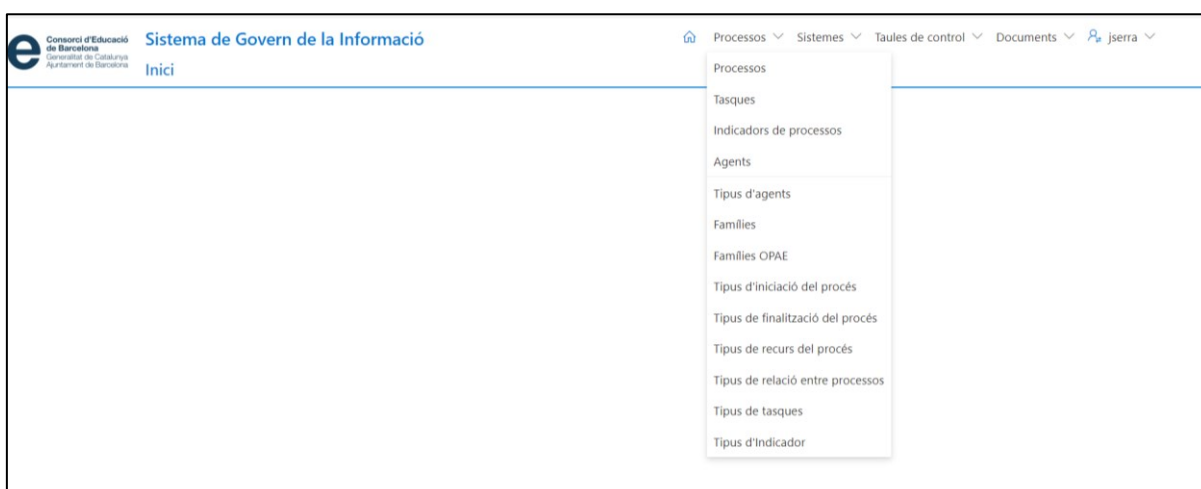
4.3.2 Funcionalitats de manteniment de dades

CATPROC 2.0 permet donar d'alta, modificar i donar de baixa totes les entitats del model de dades que es descriu a continuació. El manteniment de les entitats s'agrupa en quatre menús d'usuari.

El primer àmbit de manteniment permet gestionar els processos i totes les entitats que s'hi relacionen. Així, permet fer el manteniment dels processos, les tasques en què es descomponen, i els agents que hi participen. També permet fer el manteniment de tots els diccionaris que s'utilitzen en els atributs d'aquestes entitats: els tipus d'agent, les famílies de processos, els tipus d'iniciació i finalització dels processos, els tipus de recursos o revisions dels processos (administratius), els tipus de relacions entre processos, i els tipus de tasques. També està preparat per emmagatzemar indicadors o mètriques sobre els processos, tot i que aquesta opció no s'ha utilitzat en aquest projecte.

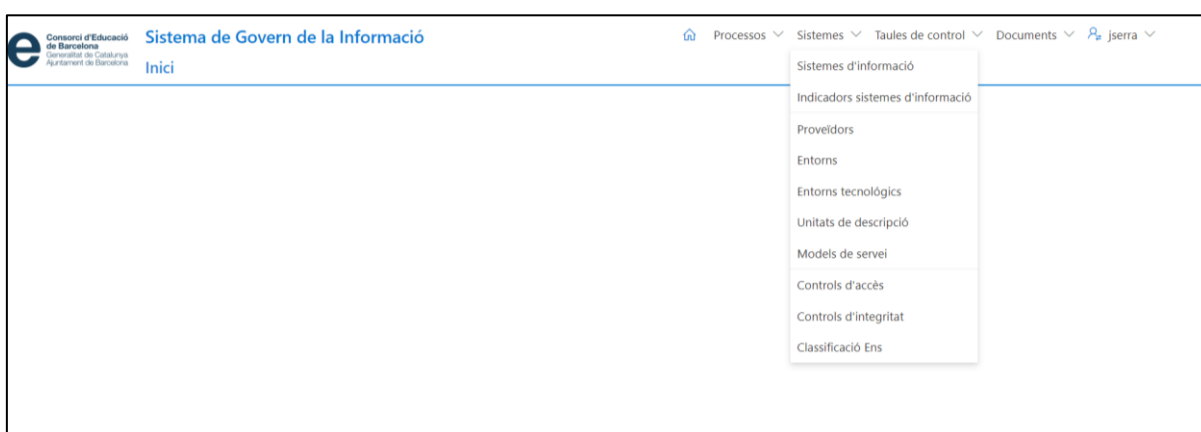
¹⁹⁶ A dia d'avui l'autor.

Figura 4-1: Menú de gestió de processos



El segon àmbit de manteniment permet gestionar els sistemes d'informació, juntament amb tots els diccionaris que apliquen al control dels seus atributs. Així doncs permet gestionar els vocabularis de proveïdors, entorns, entorns tecnològics (tecnologies), unitats de descripció, models de servei, controls d'accés, controls d'integritat i classificacions de seguretat.

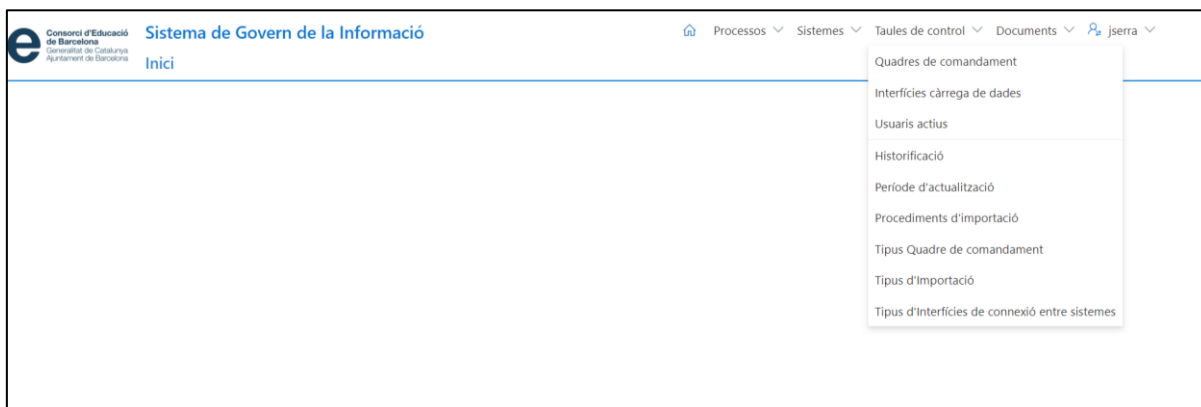
Figura 4-2: Menú de gestió de sistemes d'informació



El tercer àmbit de manteniment permet gestionar els atributs de les relacions entre sistemes, com també les visualitzacions creades a partir de les dades dels sistemes (quadres de comandament, o *dashboards*). Aquest menú es va incorporar a CATPROC 1.0 quan aquesta eina va començar a ser utilitzada de forma habitual per l'Oficina de Dades del Consorci d'Educació de Barcelona per al control de les connexions de les diferents fonts de dades a un banc de dades centralitzat, i per a l'inventari dels productes d'informació generats en forma de visualitzacions interactives. El menú permet gestionar

també el diccionaris dels diferents atributs de les relacions entre sistemes: historificacions de dades, períodes d'actualització, procediments d'importació, tipus d'importació, i tipus d'interfícies de connexió, com també els tipus de *dashboards* i els usuaris que hi accedeixen.

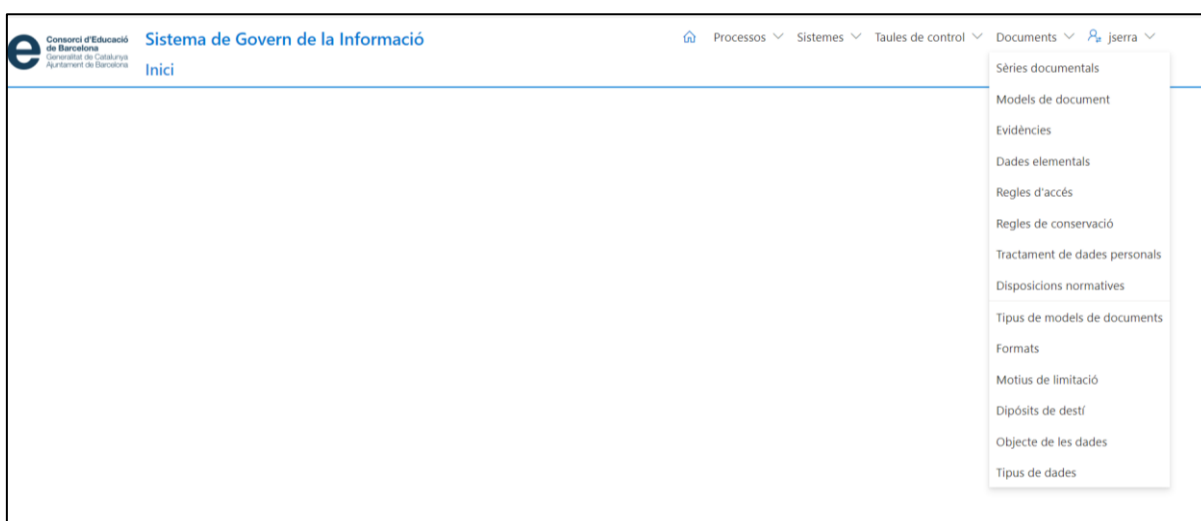
Figura 4-3: Menú de gestió d'interaccions entre sistemes d'informació



El quart àmbit és el que concentra la gestió dels diferents elements del sistema de gestió documental del Consorci. Aquests són les sèries documentals, els models de document i les evidències que componen cada sèrie documental. Permet gestionar també les regles de cicle de vida de cada sèrie documental (conservació i accés), i també el seu context jurídic (disposicions normatives) i de seguretat (tractaments de dades personals). Com en els altres casos, incorpora el manteniment dels diccionaris dels atributs d'aquestes entitats: tipus de model de document, formats, motius de limitació de les regles d'accés, i dipòsits de destí de les regles de conservació.

Aquest àmbit incorpora també els elements necessaris per completar el sistema de gestió documental amb un catàleg de dades que en permeti el seu govern integral. Així doncs trobem en aquest menú la gestió de les dades elementals, i també el manteniment dels dos atributs més rellevants del catàleg de dades: els tipus de dades, i els objectes descrits per les dades. Atesa la creixent rellevància de l'objecte descrit en l'estudi, i el seu vincle amb un possible plantejament ontològic, aquest darrer diccionari s'ha configurat en el model de dades no com un mer diccionari, sinó com una entitat independent amb els seus propis atributs.

Figura 4-4: Il·lustració 4. Menú de gestió de les entitats documentals



Cada entitat disposa d'una fitxa pròpia per al seu manteniment, a través de la qual es poden actualitzar tant les seves dades com les relacions que manté amb altres entitats del model.

Figura 4-5: Exemple de fitxa de manteniment de dades

The screenshot displays the 'Sèries documentals' maintenance form. The form contains the following fields and options:

- Codi i nom procés ***: Dropdown menu with 'BEQ001 Subvencions' selected.
- Codi ***: Text input with 'BEQ001-003'.
- Títol de la sèrie documental ***: Text input with 'Subvencions per al finançament de despeses de personal dels monitors de menjador, esb...'. Includes checkboxes for '100% digital' and 'Vital'.
- Classificació**: Text input with 'Y0715'.
- Títol segons quadre de classificació corporatiu**: Text input with 'Subvencions'.
- Caràcter ***: Dropdown menu with 'Secundària' selected.
- Tipus Productor ***: Dropdown menu with 'CEB' selected.
- Codi i nom productor ***: Dropdown menu with 'CEB-20151 Atenció i Tràmits de l'administració educativa' selected.
- criteri per donar nom als expedients**: Text input with 'N-Codi i nom de centre'.
- Tractaments de dades personals relacionat**: Dropdown menu.
- Data alta**: Date picker with '20/04/2017'.
- Data de baixa**: Date picker with '30/12/2099'.

At the bottom, there are tabs for 'Descripció', 'Dependències', 'Evidències', 'Regles d'accés', and 'Regles de conservació'. Below the tabs are four buttons: 'Guardar' (blue), 'Tornar' (grey), 'Reiniciar' (grey), and 'Eliminar' (red).

Com a mesura per garantir la integritat del conjunt, l'eina no permet la eliminació d'entitats que tinguin relacions actives amb altres entitats del model.

4.3.3 Funcionalitats d'interoperabilitat del sistema

CATPROC 2.0 incorpora diverses funcionalitats per a l'exportació de la informació recollida. Dues d'aquestes funcionalitats són transversals a tots els àmbits i entitats. La primera, accessible des de la vista general o llistat d'entitats, consisteix en l'exportació de les dades de totes les columnes visualitzades. Aquesta exportació es pot fer a format *Microsoft Excel* o *Adobe PDF*.

Figura 4-6: Opcions d'exportació de les llistes

Codi	Nom del Procés ↑	Tipus procés ↑	Tipus Responsable	Codi i nom unitat responsable	Vigent
CRD001	Coordinació de professionals d'atenció educativa	De Suport	CEB	CEB-20149 Àrea d'Educació Inclusiva	✓
CRD002	Coordinació dels Serveis d'Orientació	Clau	CEB	CEB-22614 Orientació i Acompanyament a l'Alumnat	✓
CRD003	Coordinació dels EAPs i Serveis educatius específics	Clau	CEB	CEB-20149 Àrea d'Educació Inclusiva	✓
CRD004	Coordinació tècnica i administrativa dels LICs i CRPs	Clau	CEB	CEB-20148 Capacitació Docent i Programes Educatius	✓
CRP001	Gestió i préstec de material	De Suport	CEB	CEB-20148 Capacitació Docent i Programes Educatius	✓
CRP002	Activitats de dinamització	De Suport	CEB	CEB-20148 Capacitació Docent i Programes Educatius	✓

La segona opció comuna a totes les entitats és la generació d'una fitxa descriptiva de l'entitat i de totes les seves relacions, que es genera des de la fitxa de manteniment de l'entitat en format *Adobe PDF*.

Figura 4-7: Exemple de fitxa descriptiva de l'entitat Procés

CEB. Catàleg de Processos		FITXA DEL PROCÉS		Pàg. 1 de 11
				Data 18/03/2023
DADES DEL PROCÉS				
Codi identificador del procés	INF001			
Tipus de procés	De Suport			
Macroprocés	No			
Família OPAE del procés				
Àmbit-	CEB			
Descripció del procés	<p>Objectiu: recopilació, emmagatzemament i explotació de dades que demanen els usuaris del CEB.</p> <p>Descripció: els usuaris del CEB demanen a l'USIDP informació basada en dades, que poden ser externes o generades pel propi CEB. Si ja es disposa d'aquestes dades, la USIDP atén la petició generant els indicadors que s'hagin sol·licitat, tot sol·licitant si escau el suport de l'ATIC. Si no es disposa d'aquestes dades, l'USIDP fa les gestions necessàries per aconseguir les dades i incorporar-les al banc de dades del CEB amb el suport de la ATIC. Aquest procés de captura de dades també es pot iniciar com a resultat de les accions de millora que es porten a terme des de la USIDP o l'ATIC sobre diferents àmbits de gestió del CEB, o amb motiu del desenvolupament i posada en producció de qualsevol nova solució tecnològica al CEB.</p> <p>El manteniment del banc de dades es fa amb la importació periòdica de les fonts de dades ja integrades al banc. El sistema processa aquestes dades i les prepara per a la consulta. El propi sistema conserva una còpia historificada de les dades processades. I periòdicament també s'actualitzen tots els quadres de comandament i eines de visualització de les dades del Banc de Dades.</p>			
Unitat responsable	CEB-20150 Sistemes d'Informació i disseny de processos			

A part d'aquestes dues opcions, des de la fitxa de procés, com a node principal del model de dades, existeixen dues opcions més d'exportació. La primera és la guia del procediment, una versió de la fitxa de procés orientada a obtenir una descripció textual estructurada d'aquells processos per als quals s'ha pogut definir amb detall les tasques que els componen.

Figura 4-8: Exemple de guia de procediment

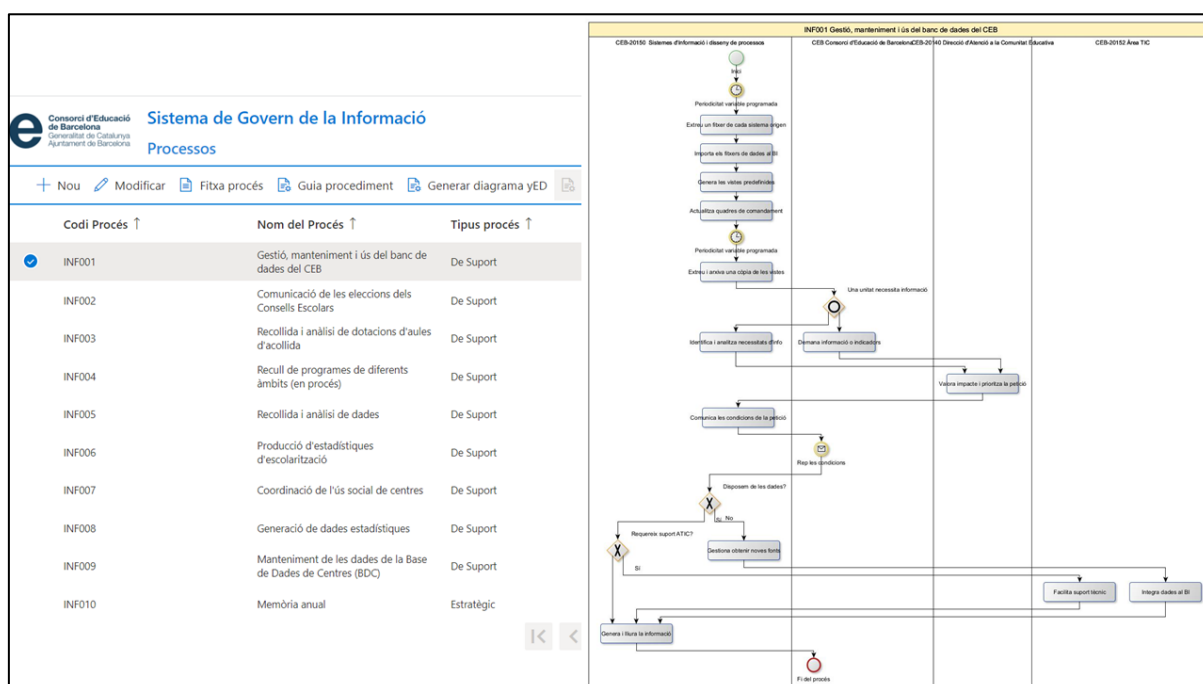
CEB. Catàleg de Processos		GUIA DE PROCEDIMENT		Pàg. 1 de 5
				Data 18/03/2023
Procés: INF001 Gestió, manteniment i ús del banc de dades del CEB				
Família del procés	Anàlisi i producció d'informació			
Descripció	<p>Objectiu: recopilació, emmagatzemament i explotació de dades que demanen els usuaris del CEB.</p> <p>Descripció: els usuaris del CEB demanen a l'USIDP informació basada en dades, que poden ser externes o generades pel propi CEB. Si ja es disposa d'aquestes dades, la USIDP atèn la petició generant els indicadors que s'hagin sol·licitat, tot sol·licitant si escau el suport de l'ATIC. Si no es disposa d'aquestes dades, l'USIDP fa les gestions necessàries per aconseguir les dades i incorporar-les al banc de dades del CEB amb el suport de la ATIC. Aquest procés de captura de dades també es pot iniciar com a resultat de les accions de millora que es porten a terme des de la USIDP o l'ATIC sobre diferents àmbits de gestió del CEB, o amb motiu del desenvolupament i posada en producció de qualsevol nova solució tecnològica al CEB.</p> <p>El manteniment del banc de dades es fa amb la importació periòdica de les fonts de dades ja integrades al banc. El sistema processa aquestes dades i les prepara per a la consulta. El propi sistema conserva una còpia historificada de les dades processades. I periòdicament també s'actualitzen tots els quadres de comandament i eines de visualització de les dades del Banc de Dades.</p>			
TASQUES DEL PROCÉS				
Seqüència	Agent / Descripció	Entrades	Sortides	
00	CEB-20150 Sistemes d'Informació i disseny de processos: Inici del procés			
01	CEB-20150 Sistemes d'Informació i disseny de processos: Amb periodicitat variable programada			
02	CEB-20150 Sistemes d'Informació i disseny de processos: La USIDP extreu de cada sistema origen les dades generades pel sistema. Segons la font de dades, l'extracció és totalment manual, semiautomàtica (amb el suport de RPA) o automàtica (per connexió directa a l'origen de les dades). Les fonts de dades d'on s'obtenen els fitxers són les que s'indiquen al Catàleg de Processos i Sistemes del CEB (apartats Sistemes d'informació i Interfícies).			

I la segona, també possible únicament per aquells processos que tenen informades totes les seves tasques, és la generació d'un fluxgrama detallat, que descriu la seqüència completa del procés. Aquest fluxgrama es genera automàticament a partir dels atributs de les tasques que componen el procés. El fluxgrama es genera en un format estàndard per a grafs i diagrames, el format *GraphML*,¹⁹⁷ i es representa utilitzant els conceptes i notació de BPMN.¹⁹⁸

¹⁹⁷ *GraphML* és un format basat en XML per a la representació de grafs i diagrames. Per a les especificacions vegeu <http://graphml.graphdrawing.org/specification.html>. Per a més informació vegeu Brandes et al. 2002.

¹⁹⁸ *Business Process Model and Notation* (BPMN), actualment en la versió 2.0.2, és l'estàndard de facto per a la representació gràfica de processos en qualsevol tipus d'organització. Des de 2005 BPMN és promogut i mantingut per *Object Management Group* (OMG), un consorci d'empreses tecnològiques. OMG va participar també en el seu desenvolupament, amb l'objectiu de disposar d'un mètode estàndard de representació de seqüències d'activitats que facilités la seva posterior implementació en eines de gestió automatitzada de processos (*Business Process Management Systems*, BPM). Aquesta implementació se simplificava per la compatibilitat de BPMN amb llenguatges executables en sistemes BPM com BPEL (*Business Process Execution Language*) o XPDLL (*XML Process Definition Language*). L'any 2013 BPMN fou adoptat com estàndard per ISO a través de la norma *ISO/IEC 19510:2013 Information technology — Object*

Figura 4-9: Exemple d'exportació de fluxograma de procediment



L'eina també permet realitzar aquest procés a la inversa. A partir d'una plantilla en format *GraphML* és possible descriure el flux complet d'un procés, diagramant cadascuna de les seves tasques, els agents que les porten a terme, i informant en el mateix diagrama tots els atributs que el model de dades preveu per a l'entitat tasca. La importació d'aquest fitxer genera automàticament dins la fitxa del procés la seqüència completa de tasques i el detall dels seus atributs. Aquesta funcionalitat ha estat un factor important en l'acceleració de les tasques d'inventari i anàlisi dels processos.

4.3.4 Funcionalitats d'exploració del sistema

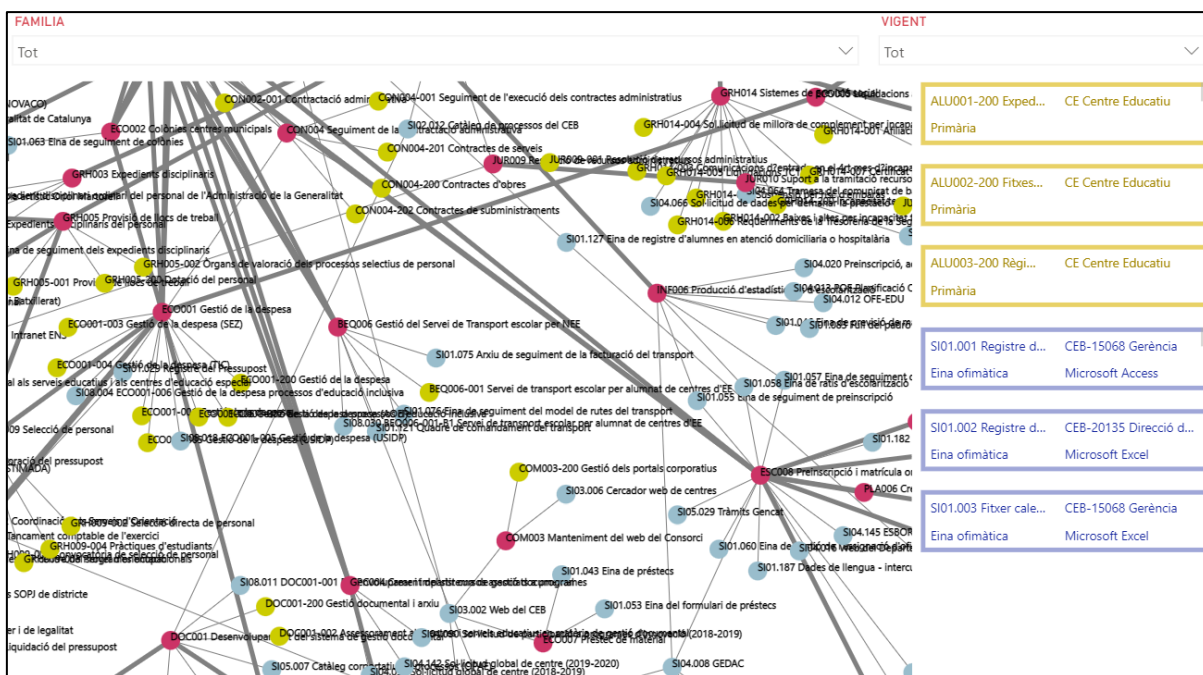
Per a l'exploració de les dades CATPROC 2.0 incorpora, dins la mateixa eina, unes funcionalitats bàsiques de cerca i filtrat de la informació, com també la possibilitat de parametritzar les visualitzacions de llista dels diferents inventaris d'entitats.

Figura 4-10: Exemple d'opcions de cerca i filtrat d'informació

The screenshot shows the 'Sistema de Govern de la Informació' interface. On the left, there is a table with columns: 'Codi i nom promotor', 'Codi', 'Nom Sistema', 'Responsable funcional', 'Entorns', and 'Entorns tecnològics'. The table lists various systems like 'CEB-20138 Direcció d'Equiptaments Educatius' and 'CEB-20139 Direcció d'Educació i Territori'. On the right, there is a 'Filtrar Sistemes d'informació' sidebar with search and filter options for 'Codi Sistema d'informació conté', 'Nom Sistema d'informació conté', 'Abast', 'Codi i nom promotor', 'Entorns', 'Entorns tecnològics', 'Unitats de descripció', 'Control d'accés', and 'Control d'integritat'. At the bottom of the sidebar are buttons for 'Cerca', 'Esborra el filtre', and 'Cancela'.

Però el gran potencial d'exploració l'ofereix la possibilitat de connectar-se directament a la base de dades de CATPROC 2.0 des d'entorns d'anàlisi i visualització de la informació. En el nostre cas, aquesta explotació mitjançant connexió directa s'ha fet preferentment utilitzant *Microsoft Power BI*.

Figura 4-11: Exemple d'entorn de visualització: mapa global de processos, sèries documentals i sistemes d'informació



4.4 Model de dades

L'eina de recollida s'ha implementat, a nivell de base de dades, sobre un model relacional. Per aquest motiu, el model de dades que es presenta a continuació s'implementa a nivell físic com un model mono-entitat, a diferència dels models més recents que s'han elaborat en l'àmbit arxivístic i que contempnen un model multi-entitat o de base ontològica.¹⁹⁹

Això no exclou que en la descripció del model lògic, i especialment en el detall dels seus components, s'hagin considerat relacions pròpies dels models multi-entitat o de base ontològica, com són les relacions d'inclusió o de dependència, i també les relacions amb atributs propis.

La terminologia per a la denominació dels components del model és difícil de determinar a partir dels referents emprats. Així, mentre un model multi-entitat, com **e-EMGDE**, utilitza els termes "Element" i "Sub-element" integrats en un mateix model, un altre model amb la mateixa orientació, com **RiC**, distingeix entre "Entitats" i "Atributs". Per aquest motiu, s'ha optat per utilitzar la terminologia relacional clàssica per a descriure el model de dades, diferenciant entre "Entitats", "Atributs" i "Relacions". El significat i utilització de cada entitat, atribut i relació es descriuen en les dues guies metodològiques del capítol 3, la Guia d'inventari i mapatge (**GMAP**) i la Guia de dadificació de processos (**GDAD**).

Pel que fa a les entitats considerades, el model que es proposa és un model orientat a descriure els elements que intervenen en el govern de la informació, no les seves concrecions en forma d'objectes documentals. Aquesta concreció és una part important de la majoria de models de descripció d'origen arxivístic, que s'orienten a descriure documents i expedients concrets. En el model que es proposa, només l'entitat **ENT004 Evidència** és equiparable a unitat o objecte documental, i només s'hi apropa des de la perspectiva tipològica, no d'instanciació.

En la descripció del model només es descriu el model lògic, no el físic. Això implica:

1. Que les entitats que es descriuen són les entitats nuclears, corresponents a coses del món real que poden ser instanciades dins el sistema de governança.
2. Que els atributs que es descriuen són atributs-concepte, elements d'informació que poden ser simples o compostos, i que poden ser validats contra diccionaris externs.
3. Que les relacions que es descriuen són relacions lògiques, de les quals només es determina el grau, obligatorietat i significació, però no el conjunt de claus per implementar-les.

¹⁹⁹ També anomenat multidimensional per RiC.

No es representen en aquest model els diccionaris de validació d'atributs (valors controlats) que es deriven de la implementació d'aquest model, i que en el model físic tenen una gestió diferenciada, llevat dels casos en què aquesta validació prové d'una llista tancada de valors.

L'esquema amb el qual es descriuen les **entitats** és el que s'indica a continuació:

Identificador	Identificador únic de l'entitat dins el model de dades. L'identificador es compon del prefix "ENT" seguit d'un número correlatiu de tres dígits.
Nom formal	És el nom amb el qual s'identifica de forma unívoca aquesta entitat dins el model.
Nom curt	És una forma compacta del nom formal, als efectes de facilitar l'esment d'aquesta entitat en diagrames, representacions gràfiques o descripcions textuales.
Domini	És la descripció dels elements del món real que seran descrits per aquesta entitat. Pot incloure comentaris explicatius per facilitar la comprensió de l'entitat.
Atributs	Detall dels identificadors i noms formals de tots els atributs que caracteritzen aquesta entitat, i que seran descrits a continuació d'aquesta taula.
Relacions	<p>Detall de les relacions d'aquesta entitat amb altres entitats del model, indicant en forma narrada les característiques de la relació. Així, s'utilitzarà un verb per indicar el sentit de la relació, que estarà:</p> <ul style="list-style-type: none"> • En present d'indicatiu, quan la relació sigui obligatòria (sempre) • En infinitiu precedit de "Pot", quan la relació sigui opcional (no sempre) <p>A continuació, s'indicarà la cardinalitat de la relació, en funció de si la relació és amb "un" (1) sol element de l'entitat destí, o amb "un o més" (N) elements de l'entitat destí.</p> <p>Finalment, s'indica sempre en negreta l'entitat destinatària de la relació.</p>
Fonts	Se citen en aquest apartat les normes de descripció que fan referència a aquesta entitat, i a través de les quals s'ha identificat la necessitat de disposar d'aquesta entitat en el nostre model. Les normes se citen per la seva abreviatura (indicada a l'apartat "Models de referència i abreviatures per citació"), i es marquen en negreta. S'inclouen quan escau aclariments sobre el significat d'aquella entitat en la font citada.

A continuació de cada entitat es descriuen, un per un, els **atributs** que la conformen d'acord amb el següent esquema:

Identificador	Identificador únic de l'atribut dins el model de dades. L'identificador es compon del l'identificador complet de l'entitat, seguit d'un guió i d'un número correlatiu de tres dígits.
----------------------	---

Nom formal	És el nom amb el qual s'identifica de forma unívoca aquest atribut.
Nom curt	És una forma compacta del nom formal, als efectes de facilitar la utilització d'aquest atribut en interfícies, models físics de dades o estructures de base de dades.
Domini	És la descripció detallada de la característica de l'entitat que es recollirà en aquest atribut.
Tipus de dades	<p>Tipus de dades que es podran informar en aquest atribut. El tipus de dades en el model lògic obeeix bàsicament als tipus de validació de les quals podran ser objecte les dades de l'atribut en les seves implementacions físiques.</p> <p>Els tipus de dades contemplats són els següents:</p> <ul style="list-style-type: none"> • Alfanumèric • Alfanumèric descriptiu • Alfanumèric categòric • Numèric • Data • Booleà (sí/no) • Usuari (vinclat a directori d'usuaris)
Mín. ocurrencies	<p>Quantitat mínima de valors que pot prendre el camp. Aquest element pot prendre dos valors: 0 o 1</p> <p>Aquest element indica el caràcter obligatori del camp.</p>
Màx. ocurrencies	<p>Quantitat màxima de valors que pot prendre el camp. Aquest element pot prendre tres valors:</p> <ul style="list-style-type: none"> • 1 • Qualsevol número enter • N (número indeterminat de valors) <p>Aquest element indica el caràcter repetible del camp.</p>
Regles de validació	<p>Qualsevol regla que determini el valor, forma o estructura d'aquest atribut. Les regles poden ser:</p> <ul style="list-style-type: none"> • Llistes de valors controlats • Fonts externes de valors controlats • Formats específics de configuració de l'atribut • Requisits de validesa semàntica de l'atribut
Valor per defecte	Valor per defecte que ha de prendre l'atribut, en cas que n'hagi de prendre algun. En cas contrari, s'indica amb "No aplica".
Fonts	Se citen en aquest apartat les normes de descripció que fan esment d'aquest atribut, estigui vinculat o no amb l'entitat descrita. En cada cas se cita la norma de descripció per la seva abreviatura (indicada a l'apartat "Models de referència i abreviatures per citació") i l'identificador i nom de l'atribut equivalent en aquella norma, amb els aclariments que siguin pertinents. Tant l'entitat com l'atribut es marquen en negreta.
Exemples	Es facilita un exemple d'una instància concreta d'aquest atribut.

Aquest model de dades s'ha utilitzat no només per a la recollida d'informació per aquesta tesi, sinó també com a eina de governança de la informació al Consorci d'Educació de Barcelona. Aquest ha estat un ús derivat d'haver observat progressivament el potencial que mostrava com a instrument de governança i control, més enllà de l'original potencial d'anàlisi. Per aquest motiu, s'ha afegit al model alguns atributs que provenen bàsicament d'aquesta necessitat operativa, i que poden tenir utilitat per altres organitzacions similars al Consorci, però que no s'utilitzen en aquesta tesi per a finalitats analítiques. En la descripció del domini de cada atribut s'indica aquesta condició.

Finalment es completa el model amb la descripció de les relacions entre les diferents entitats. Cada relació es descriu indicant també les característiques i significat de la relació si es considera en sentit invers. La descripció de cada relació segueix el següent model, i es completa, si escau, amb la descripció dels atributs propis d'aquella relació:

Identificador	Identificador únic de la relació dins el model de dades. L'identificador es compon del prefix "REL" seguit d'un número correlatiu de tres dígits.
Domini	S'indica l'entitat del model a la qual s'origina la relació.
Rang	S'indica l'entitat del model a la qual finalitza la relació.
Significat de la relació	S'indica l'acció a la qual correspon la relació entre el domini i el rang.
Relació inversa	S'indica l'acció a la qual correspon la relació si es considera en sentit invers.
Obligatorietat	Es fa constar si la relació és obligatòria o opcional. En cas que la obligatorietat no apliqui en ambdós sentits, s'indica el nivell d'obligatorietat de cadascun dels sentits de la relació, separat per "/".
Cardinalitat	S'indica el grau de la relació (1 o N) en el sentit domini-rang.
Definició	És la descripció detallada del significat de la relació, juntament amb el context que permet entendre l'obligatorietat i el grau de la relació. S'inclouen quan escau aclariments complementaris.
Atributs	En cas que la relació es caracteritzi amb atributs propis, més enllà dels que siguin implementables amb claus externes o taules d'intersecció, s'indica en aquest apartat aquests atributs, i es descriuen a continuació d'aquesta fitxa. En cas contrari s'indica "Sense atributs propis".
Fonts	Se citen en aquest apartat les fonts que fan esment d'aquesta relació, tant si en la norma origen estan contemplades com a relació, com si estan contemplades com entitats o atributs d'entitat. En cada cas se cita la norma per la seva abreviatura (indicada a l'apartat "Models de referència i abreviatures per citació"), i a continuació l'identificador i nom de l'element equivalent, amb els aclariments que siguin pertinents. Tant l'entitat com qualsevol altre element citat es marquen en negreta.

4.5 Models de referència i abreviatures per citació

L'objectiu inicial d'aquest model era estructurar la recollida d'informació per tal de construir un mapa de tots els elements vinculats a la gestió de la informació en el cas d'estudi, el Consorci d'Educació de Barcelona. I fer-ho partint des de la perspectiva de l'arxivística i la gestió documental, per poder valorar l'adequació dels seus models i conceptes al tractament d'actius d'informació en un entorn fortament orientat a dades. Per a la definició d'aquest model s'ha tingut en compte les principals normes de descripció vigents en l'àmbit arxivístic, juntament amb una selecció de models i esquemes de metadades sense un origen pròpiament arxivístic, però aplicables també a la descripció d'actius d'informació (i especialment de conjunts de dades).

En paral·lel a aquest objectiu inicial, l'ús d'aquest model per a la recollida d'informació ha posat de relleu també el seu potencial per a la construcció d'un sistema de governança de la informació, i especialment de la dada, el qual constituïa un dels objectius secundaris d'aquesta tesi. Aquesta segona utilització, recurrent en el desenvolupament del projecte, ha fet incorporar al model atributs molt orientats a la seva implementació pràctica al Consorci d'Educació de Barcelona i entitats de característiques similars (i per tant poc presents als models de referència), i especialment d'entitats inexistents en la majoria de models consultats però que considerem rellevants per a la governança de la informació, com per exemple l'entitat **ENT006 Dada elemental**, **ENT011 Sistema d'informació o aplicació**, **ENT013 Visualització de dades** o **ENT014 Objecte** (descriu).

Un segon factor de desencaix ha estat que la majoria de models consultats estan dissenyats per a la creació d'instruments de descripció, no de governança, i per aquest motiu bona part d'elements descriptius dels que aporten només tenen sentit sobre instanciacions concretes. Això es fa especialment palès amb l'entitat **ENT004 Evidència documental**, que en aquest model es concep com una entitat abstracta (document tipus que apareix en un circuit documental), i en la majoria de models consultats es concep com una instància física (un document existent en el món real). Per aquest motiu, en nombrosos casos s'ha hagut de forçar una correspondència entre entitats i atributs de les diferents normes que, malgrat que existent, només ho és en la capa conceptual i no en la implementació d'aquella entitat.²⁰⁰

²⁰⁰ En el cas de l'entitat **ENT004 Evidència documental** s'ha assimilat majoritàriament al concepte de "document" o "*record*", tot i que en el nostre model no es representa en cap cas la seva instanciació física. La nostra proposta té punts de contacte amb el concepte de "mapa documental" proposat per Carlota Bustelo i Elisa García-Morales (Zapata, 2002; García-Morales, 2014).

Els models buidats es poden agrupar en dues grans famílies (Niu, 2013):²⁰¹

1. Els models originaris de l'àmbit arxivístic, i orientats preferentment a facilitar una descripció que sigui útil tant per als arxius històrics com per als arxius administratius. Aquests models inclouen les normes emanades del Consell Internacional d'Arxius, bàsicament la família ISAD(G) (**ISAD**, **ISAAR**, **ISDF** i **ISDIAH**) i la nova norma *Records in Contexts* (**RIC**), i també les normes de la *Society of American Archivists* (**EAD** i **EAC**) i del *Ministerio de Cultura i Deporte* (**NEDA**).²⁰²
2. Els models originaris de la gestió documental, i especialment pensats per a la gestió documental electrònica (Cottin & Dessolin, 2012). Són els models que provenen del món de la norma ISO 15489 de gestió documental, i específicament del desplegament del seu esquema de metadades, la norma **ISO 23081**. La tradició australiana del *records management* (McKemmish et al., 1999), gènesi de la norma ISO 15489, es fa present amb la norma **AGRkMS**, en altres propostes com la norma espanyola **e-EMGDE**, i també amb un dels primers intents de conjuminar les normes arxivístiques amb la descripció de conjunts de dades, la norma **ONDC**.

Aquest conjunt s'ha combinat amb una selecció de models d'altres àmbits. Aquests són la descripció de dades i la creació de catàlegs de dades, amb l'ontologia **DCAT** vinculada a **PROV**, la preservació digital, amb el model **SIARD**, i un model singular que, tot i emanat de l'àmbit de la gestió documental, té molts pocs vincles amb la família ISO 23081, el model de **MoReq** (Henttonen, 2009; Vieira et al., 2012).²⁰³ No s'han considerat altres models, com el *DDI-Codebook* de la *Data Documentation Initiative Alliance* (Rasmussen & Blank, 2007; Vardigan et al., 2008; Williams et al., 2017),²⁰⁴ per limitar el buidatge a un volum adequat al seu objectiu inicial de creació d'un instrument per a la recollida d'informació per al projecte.²⁰⁵

²⁰¹ Se citen les normes utilitzant l'abreviatura que s'indica en aquest mateix apartat 4.5.

²⁰² Conceptualment s'inclouen aquí totes les normes derivades i versions locals, com són, en el cas català, la Norma de descripció arxivística de Catalunya (NODAC), els Vocabularis de metadades del Grup d'Innovació Tecnològica del Departament de Cultura i Mitjans de comunicació, o l'Instrument tècnic Esquema de Metadades e-SIGeDA. V.1.0.

²⁰³ *MoReq2010* és un model elaborat directament per la indústria del *software* de gestió documental, el qual, tot i molt ajustat a les necessitats de la gestió documental electrònica, té una orientació força diferent a la de la norma ISO 23081, molt influenciada encara per l'arxivística històrica.

²⁰⁴ La *Data Documentation Initiative Alliance* és un projecte internacional que té per objectiu desenvolupar estàndards de metadades i productes semàntics per descriure conjunts de dades en ciències socials. *DDI-Codebook* és una versió reduïda de l'estàndard en XML, orientada a documentar dades de recerca.

²⁰⁵ S'ha descartat models de metadades com *PREMIS* (<https://www.loc.gov/standards/premis/>) o *Dublin Core* (<https://www.dublincore.org/>) amb la voluntat de limitar-nos als models d'origen

La principal conclusió del buidatge d'aquests models és la pervivència d'una sèrie de conceptes, propis de l'arxivística històrica i dels inicis de la gestió documental, que han canviat molt poc amb la progressiva implantació de la tecnologia digital, i que segueixen fortament orientats a la descripció d'objectes (Niu, 2013). També s'ha observat una coincidència molt alta entre els elements descriptius de la majoria de models, probablement producte d'aquesta pervivència dels conceptes clàssics. Evidentment els models han evolucionat en la seva estructura, preferentment cap a plantejaments ontològics (Gueguen et al., 2013; Llanes-Padrón & Pastor-Sánchez, 2017), però l'evolució no s'ha traslladat a la capa conceptual, llevat de l'àmbit de la preservació digital, on sí que s'ha evolucionat de la mà del model OAIS (Niu, 2013). Prova d'això és la minsa presència del concepte de sistema d'informació o aplicació com un element amb entitat pròpia dins el sistema de gestió, el qual tots els models redueixen a un tipus d'agent, o la poca reflexió sobre l'acció de classificar com a expressió de relacions entre entitats, que molts models encara redueixen a un atribut descriptiu (codi de classificació).

Els models utilitzats han estat els següents.²⁰⁶

Taula 4-1: Models de referència per a l'elaboració del model de dades

Abreviatura	Referència ²⁰⁷
AGRkMS	<i>Australian Government Recordkeeping Metadata Standard (AGRkMS)</i> . Versió 2.2. [en línia], 2015. Canberra: National Archives of Australia. Disponible a: https://www.naa.gov.au/sites/default/files/2019-09/AGRkMS-Version-2.2-June-2015_tcm16-93990_1.pdf .
DCAT	<i>Data Catalog Vocabulary (DCAT)</i> . Versió 2. [en línia], 2020. S.l.: W3C Recommendation. Disponible a: https://www.w3.org/TR/vocab-dcat-2/ .
EAC	<i>Encoded Archival Context - Corporate Bodies, Persons, and Families (EAC-CPF)</i> . Versió 2.0. [en línia], 2022. S.l.: Technical Subcommittee for Encoded Archival Standards, Society of American Archivists. Disponible a: https://eac.staatsbibliothek-berlin.de/ .

arxivístic, però s'ha incorporat elements d'ambdós models a nivell conceptual (en el cas de *PREMIS*) i a nivell ontològic (via *PROV-O*).

²⁰⁶ En aquesta relació i en la descripció que segueix se citen per ordre alfabètic de l'abreviatura utilitzada. En l'apartat de "Fonts" de les descripcions de les entitats, atributs i relacions, l'ordre s'ha disposat per proximitat dels conceptes o per ordre cronològic de l'aparició de cada norma o model de descripció.

²⁰⁷ Per homogeneïtzar i fer comparable la terminologia, totes les normes de descripció se citen i referencien en la seva versió en llengua anglesa, llevat d'aquelles que estan publicades originàriament en castellà.

Per a la citació bibliogràfica de les normes de descripció, a diferència de la bibliografia final, s'ha considerat més aclaridor utilitzar el format de citació ISO 690.

Abreviatura	Referència ²⁰⁷
EAD	<i>Encoded Archival Version EAD3 1.1.1</i> [en línia], 2019. Chicago: Technical Subcommittee for Encoded Archival Standards, Society of American Archivists. Disponible a: https://www.loc.gov/ead/ .
e-EMGDE	<i>Esquema de Metadatos para la Gestión del Documento Electrónico (e-EMGDE)</i> . Versió 2.0. Documentació complementaria a la Norma Tècnica de Política de gestió de documents electrònics. [en línia], 2016. Madrid: Direcció de Tecnologies de la Informació i les Comunicacions, Ministeri de Hacienda y Administraciones Pùblicas. Disponible a: https://administracionelectronica.gob.es/pae/Home/dam/jcr:26d90313-7af0-4f69-9d62-c853d55d7627/Esquema_Metadatos_e-EMGDE_2016.pdf .
ISAAR	<i>ISAAR(CPF): International standard archival authority record for corporate bodies, persons and families</i> [en línia], 2004. 2nd ed. Paris: International Council on Archives. Disponible a: https://www.ica.org/sites/default/files/CBPS_Guidelines_ISAAR_Second-edition_EN.pdf .
ISAD	<i>ISAD(G): general international standard archival description</i> [en línia], 2000. 2nd ed. Ottawa: International Council on Archives. Disponible a: https://www.ica.org/sites/default/files/CBPS_2000_Guidelines_ISAD%28G%29_Second-edition_EN.pdf .
ISDF	<i>ISDF: International Standard for Describing Functions</i> [en línia], 2007. 1st ed. Paris: International Council on Archives. Disponible a: https://www.ica.org/sites/default/files/CBPS_2007_Guidelines_ISDF_First-edition_EN.pdf .
ISDIAH	<i>ISDIAH: International standard for describing institutions with archival holdings</i> [en línia], 2008. 1st ed. Londres: International Council on Archives. Disponible a: https://www.ica.org/sites/default/files/CBPS_2008_Guidelines_ISDIAH_First-edition_EN.pdf .
ISO 23081	<p><i>ISO 23081-1:2017 Information and documentation — Records management processes — Metadata for records — Part 1: Principles</i> [en línia], 2017. S.I.: International Organization for Standardization. Disponible a: https://www.iso.org/standard/73172.html.</p> <p><i>ISO 23081-2:2021 Information and documentation — Metadata for managing records — Part 2: Conceptual and implementation issues</i> [en línia], 2021. S.I.: International Organization for Standardization. Disponible a: https://www.iso.org/standard/81600.html.</p> <p><i>ISO/TR 23081-3:2011 Information and documentation — Managing metadata for records — Part 3: Self-assessment method</i> [en línia], 2011. S.I.: International Organization for Standardization. Disponible a: https://www.iso.org/standard/57121.html.</p>

Abreviatura	Referència ²⁰⁷
MoReq	<i>MoReq2010: Modular Requirements for Records Systems</i> . Versió 1.1. [en línia], 2010. S.l.: DLM Forum Foundation. Disponible a: https://www.moreq.info/files/moreq2010_vol1_v1_1_en.pdf .
NEDA	<p><i>NEDA-MC. Modelo conceptual de descripción archivística: entidades, relaciones y atributos</i> [en línia], 2017. Madrid: Comisión de Normas Españolas de Descripción Archivística, Ministerio de Cultura y Deporte. Disponible a: https://www.culturaydeporte.gob.es/cultura/areas/archivos/mc/cneda/documentacion/normas/neda-mc.html.</p> <p><i>NEDA-Req. Datos básicos para describir documentos de archivo, agentes y funciones</i> [en línia], 2018. Madrid: Comisión de Normas Españolas de Descripción Archivística, Ministerio de Cultura y Deporte. Disponible a: https://www.culturaydeporte.gob.es/cultura/areas/archivos/mc/cneda/documentacion/normas/neda-req.html.</p> <p><i>NEDA-Voc. Terminología y glosario técnico de descripción archivística</i> [en línia], 2017. Madrid: Comisión de Normas Españolas de Descripción Archivística, Ministerio de Cultura y Deporte. Disponible a: https://www.culturaydeporte.gob.es/cultura/areas/archivos/mc/cneda/documentacion/normas/neda-voc.html.</p>
ONDC	<i>ONDC Metadata Attributes</i> [en línia], 2023. S.l.: Office of the National Data Commissioner, Australian Government. Disponible a: https://www.datacommissioner.gov.au/sites/default/files/2023-04/ONDC%20Metadata%20Attributes%202023_0.pdf .
PROV	<i>PROV-O: The PROV Ontology</i> [en línia], 2013. W3C Recommendation. Disponible a: https://www.w3.org/TR/prov-o/ .
RiC	<i>Records in contexts. Conceptual model</i> [en línia], 2021. Consultation Draft V0.2. Paris: Expert Group on Archival Description, International Council on Archives. Disponible a: https://www.ica.org/sites/default/files/ric-cm-02_july2021_0.pdf .
SIARD	<i>SIARD-2.2 Format Specification</i> (versió 2.2). [en línia], 2021. Berna: DILCIS Board, Swiss Federal Archives. Disponible a: https://siard.dilcis.eu/SIARD%202.2/SIARD%202.2.pdf .

Descrivim breument a continuació les normes de descripció utilitzades:

AGRKMS

Australian Government Recordkeeping Metadata Standard és l'estàndard desenvolupat pels Arxius Nacionals d'Austràlia per a la gestió documental en el sector públic. Defineix la informació que ha de ser capturada pels sistemes de gestió del govern d'Austràlia, especialment pels sistemes de gestió de documents electrònics, i s'adapta al model

procedimental de la norma ISO 15489, i pel que fa a les metadades al model conceptual de la norma ISO 23081.

DCAT

Data Catalog Vocabulary és un vocabulari en RDF per a la descripció de conjunts i serveis de dades, desenvolupat per facilitar la interoperabilitat entre catàlegs de dades publicats al Web. Inicialment fou desenvolupat pel *Digital Enterprise Research Institute* (DERI), un centre de recerca de la Universitat de Galway (Irlanda).

EAC

Encoded Archival Context for Corporate Bodies, Persons, and Families és un estàndard basat en XML per a la descripció i control d'autoritats preferentment en instruments de descripció d'arxiu. L'any 2011 fou adoptat com estàndard per la *Society of American Archivists*, i el 2015 els subcomitès tècnics de les normes EAD i EAC foren fusionats en el *Technical Subcommittee on Encoded Archival Standards* (TS-EAS), responsable a partir d'aquell moment del manteniment conjunt d'ambdós estàndards.

EAD

Encoded Archival Description és un estàndard basat en XML per a la codificació d'instruments de descripció d'arxiu. Fou desenvolupat pel *Technical Subcommittee for Encoded Archival Standards* de la *Society of American Archivists*, en col·laboració amb la *Library of Congress*.

e-EMGDE

La *Norma Técnica de Interoperabilidad de Política de gestión de documentos electrónicos* és la norma que, dins l'*Esquema Nacional de Interoperabilidad* (ENI),²⁰⁸ determina els requeriments que han de complir els sistemes de gestió de documents electrònics de l'Administració i sector públics. Aquesta norma inclou un detallat esquema

²⁰⁸ L'*Esquema Nacional de Interoperabilidad* fou creat pel *Real Decreto 4/2010, de 8 de enero, por el que se regula el Esquema Nacional de Interoperabilidad en el ámbito de la Administración Electrónica*, com a àmbit de compartició de tecnologia entre totes les Administracions públiques espanyoles, depenent del Ministeri competent en Administracions públiques. El nucli d'aquesta iniciativa és un conjunt de normes tècniques, les *Normas Técnicas de Interoperabilidad* (NTI), que es van revisant i adaptant a les necessitats del sector públic. La *Norma Técnica de Interoperabilidad de Política de gestión de documentos electrónicos* fou aprovada amb la *Resolución de 28 de junio de 2012, de la Secretaría de Estado de Administraciones Públicas, por la que se aprueba la Norma Técnica de Interoperabilidad de Política de gestión de documentos electrónicos*.

de metadades per a la gestió del document electrònic, inspirat en la versió 2.2 de l'*Australian Government Recordkeeping Metadata Standard*, i compatible amb la resta de normes de l'ENI.

ISAAR

La norma ISAAR(CPF) és l'estàndard per a la descripció d'entitats productores de documents d'arxiu ("*corporate bodies, persons and families*") del Consell Internacional d'Arxius. Fou desenvolupada per la *Commission on Descriptive Standards* (ICA/DDS) de forma coetània amb la norma ISAD(G), de la qual n'és complementària.

ISAD

La norma ISAD(G) és l'estàndard de descripció de documents d'arxiu del Consell Internacional d'Arxius. Fou desenvolupada per la *Commission on Descriptive Standards* (ICA/DDS), la qual ha portat a terme la revisió i edició de la segona versió de la norma.

ISDF

La norma ISDF és l'estàndard per a la descripció de funcions i activitats del Consell Internacional d'Arxius. Té el seu origen en la constatació, per part de la comunitat professional, de la manca d'homogeneïtat i sistemàtica en la identificació de funcions durant l'elaboració de les classificacions funcionals requerides per la norma ISO 15489 de gestió documental. Fou desenvolupada pel *Committee on Best Practices and Standards* (ICA/CBPS) del Consell Internacional d'Arxius.

ISDIAH

La norma ISDIAH és l'estàndard per a la descripció d'institucions que custodien fons d'arxiu. Fou desenvolupada pel *Committee of Best Practices and Standards* (ICA/CBPS) com a complement de la norma ISAD(G).

ISO 23081

La norma ISO 23081 defineix el conjunt de metadades que desenvolupa els requeriments de la norma ISO 15489. S'estructura en tres parts: la part 1, que presenta els principis, la part 2, que tracta dels reptes conceptuals i d'implementació, i la part 3, que proposa un mètode d'autoavaluació. La ISO 23081 no és estrictament un esquema de metadades, atès que no proposa cap sintaxi concreta, sinó elements de metadades

a alt nivell que agrupa en sis tipus: *identity, description, use, event plan, event history i relation*.

MoReq

Modular requirements for records systems (MoReq 2010) és un conjunt d'especificacions per a la definició de sistemes de gestió documental electrònica, que inclou un model de metadades propi. *MoReq 2010* és el resultat del *DLM Forum*, un projecte internacional promogut i finançat per la Comissió Europea des de l'any 1994.²⁰⁹ La primera versió de *MoReq*, inicials de *Model Requirements for the Management of Electronic Records*, es va publicar l'any 2001 com a resultat del tercer *DLM Forum* celebrat a Barcelona. L'any 2008 es va publicar la segona versió, *MoReq2*, i l'any 2011 es va publicar la tercera i actualment vigent versió, *MoReq2010*. En tant que model de requeriments funcionals en sistemes orientats a objectes, *MoReq2010* no descriu només les entitats i atributs (en diu metadades) de cada entitat, sinó també els seus comportaments, que anomena funcions.

NEDA

Les *Normas Españolas de Descripción Archivística* són un conjunt d'estàndards i recomanacions elaborades per la *Comisión de Normas Españolas de Descripción Archivística* (CNEDA), un grup de treball tècnic del *Ministerio de Cultura y Deporte*. Les normes es componen d'un model conceptual a alt nivell (NEDA-MC), un vocabulari tècnic (NEDA-Voc) i un conjunt de requeriments per a la descripció arxivística (NEDA-Req), que es complementen amb extensions puntuals per a determinades tipologies documentals.

ONDC

ONDC Metadata Attributes és un esquema de metadades desenvolupat per l'*Office of the National Data Commissioner* del Govern d'Austràlia. Es deriva de la publicació de la *Data Availability and Transparency Act 2022*, que estableix un marc d'interoperabilitat per a la catalogació, publicació i compartició de dades a l'administració australiana, anomenat *DATA Scheme*. Un dels elements del *DATA Scheme* és la creació d'un registre d'actius d'informació, per a la confecció del qual s'ha publicat l'esquema de metadades de l'ONDC. Conjuntament amb els Arxius Nacionals d'Austràlia, ONDC ha

²⁰⁹ El *DLM Forum* és el resultat de les conclusions del Consell Europeu del 17 de juny de 1994 el qual va acordar una major cooperació a escala europea en l'àmbit dels arxius. La primera trobada del *DLM Forum* va tenir lloc l'any 1996 a Brussel·les, i s'ha continuat celebrant cada tres anys a diferents ciutats europees. Les seves activitats es poden consultar a <https://www.dlmforum.eu/>.

mapat aquest esquema amb l'*Australian Government Recordkeeping Metadata Standard* (AGRkMS).

PROV

PROV-O és una ontologia del W3C que proporciona un conjunt de classes, propietats i restriccions que es poden utilitzar per representar i intercanviar informació de procedència generada en diferents sistemes i en diferents contextos. El seu plantejament és genèric i extensible, i no es limita a cap àmbit o disciplina en concret. Però el seu potencial per a descriure contextos arxivístics és notable, i probablement encara no explotat.

RIC

Records in Contexts és la norma de descripció arxivística desenvolupada pel Consell Internacional d'Arxius amb la voluntat d'integrar i substituir els actuals estàndards de descripció de documents (ISAD(G)), autoritats (ISAAR(CPF)), funcions (ISDF) i institucions d'arxiu (ISDIAH). La norma RiC es compon d'un model conceptual (RIC-CM), i d'una ontologia (RIC-O) codificada en OWL.²¹⁰

SIARD

SIARD correspon a l'especificació d'un format per a la preservació a llarg termini de bases de dades relacionals. Les inicials corresponen a *Software Independent Archival of Relational Databases*. La primera versió fou desenvolupada pels Arxius Federals Suïssos. La versió actual ha estat desenvolupada pel *DILCIS Board*²¹¹ en el curs del projecte E-ARK3.²¹²

²¹⁰ *Ontology Web Language*, l'estàndard del W3C per a la creació d'ontologies, disponible a <https://www.w3.org/2007/09/OWL-Overview-es.html>.

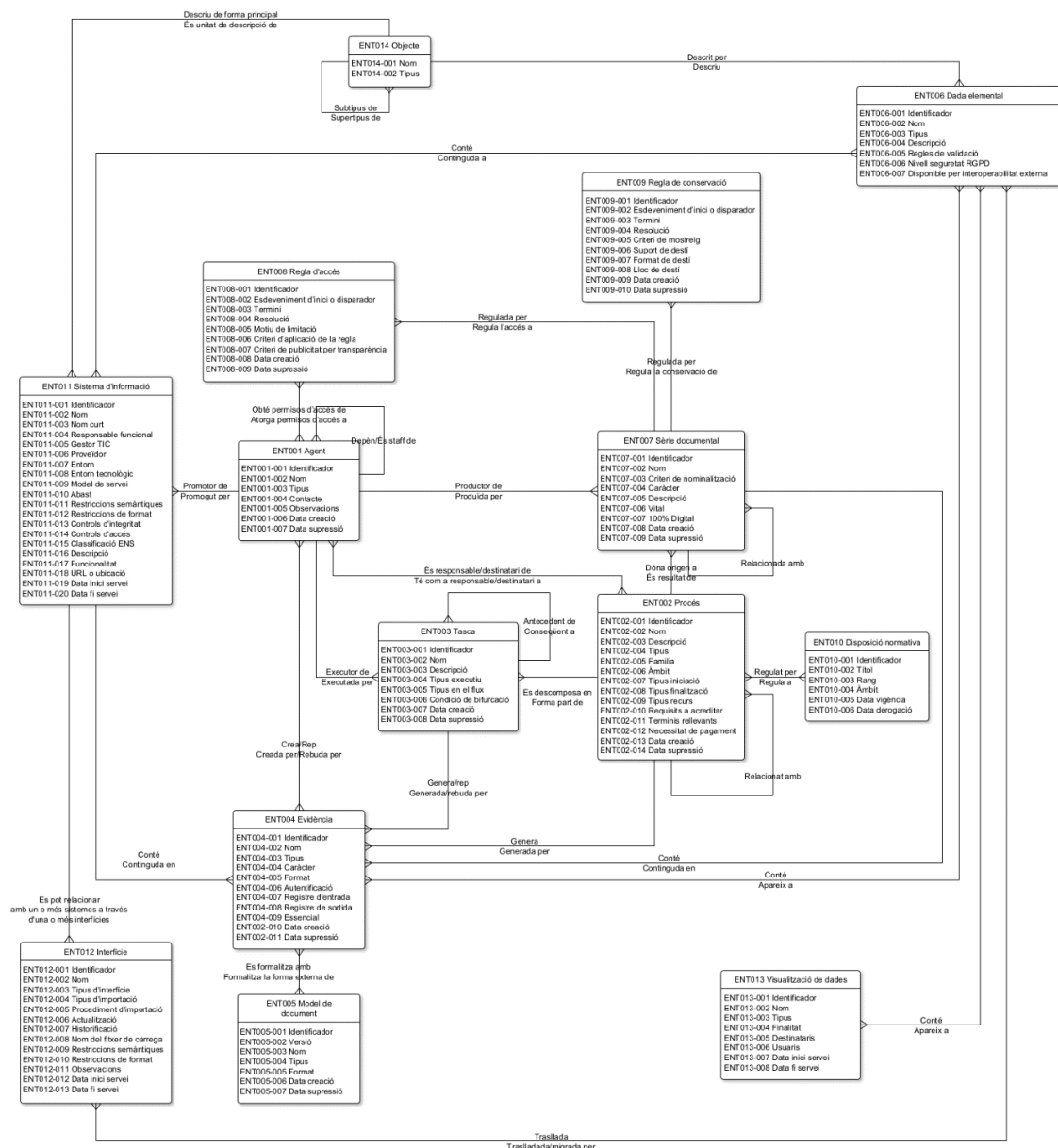
²¹¹ El *Digital Information LifeCycle Interoperability Standards Board (DILCIS Board)* és un grup d'experts internacional que té per objectiu mantenir les especificacions desenvolupades en el curs del projecte E-ARK3 (<https://dilcis.eu/>).

²¹² E-ARK (*European Archival Records and Knowledge Preservation*) fou un projecte internacional, finançat per la Comissió Europea, i desenvolupat entre 2014 i 2017, per al desenvolupament i millora de solucions i estàndards de preservació digital (<https://www.eark-project.com/>).

4.6 Descripció dels components

Descrivim a continuació el model de dades que incorpora el sistema desenvolupat. Correspon a un model entitat-relació pur, per tal d'adaptar-lo al sistema de base de dades relacional en el qual s'ha implementat (*Microsoft SQL Server*). Representem únicament el model lògic, i ho fem utilitzant la notació bàsica del diagrama de classes UML,²¹³ però indicant únicament el grau de la relació.

Figura 4-12: Diagrama entitat-relació del model de dades de l'eina CATPROC (model lògic)



²¹³ Llenguatge de modelat Unified Modeling Language versió 2.5.1 (2017), reconegut com estàndard ISO/IEC 19505-1/2:2012 Information technology — Object Management Group Unified Modeling Language (OMG UML). Disponible a <https://www.uml.org/index.htm>.

4.6.1 Entitats

4.6.1.1 Entitat: Agent

4.6.1.1.1 Descripció de l'entitat

Identificador	ENT001
Nom formal	Agent
Nom curt	Agent
Domini	<p>Es considera agent qualsevol ens que pot intervenir activament o passiva en un flux d'informació. S'entén per participació activa la creació i tramesa d'informació, i per participació passiva la recepció i arxiu de la informació.</p> <p>Els agents poden ser humans o màquines. En l'àmbit d'aquesta entitat, només es consideren els agents humans; els agents màquines es recullen a l'entitat ENT011 Sistemes d'informació.</p> <p>Els agents humans poden ser persones físiques (individus), persones jurídiques (conjunts d'individus amb existència simbòlica comuna) o bé rols (persones físiques o jurídiques tipus, identificades a partir de l'abstracció de les seves característiques comunes).</p>
Atributs	ENT001-001 Identificador ENT001-002 Nom ENT001-003 Tipus ENT001-004 Contacte ENT001-005 Observacions ENT001-006 Data creació ENT001-007 Data supressió
Relacions	REL001 Pot dependre/ser staff de un o més agents REL002 Pot ser responsable/destinatari d'un o més processos REL003 Pot ser executor d'una o més tasques REL004 Pot ser productor d'una o més sèries documentals REL005 Pot obtenir permisos d'accés d'una o més regles d'accés REL006 Pot ser promotor d'un o més sistemes d'informació REL007 Pot ser creador/receptor d'una o més evidències
Fonts	<p>Els agents són l'objecte de descripció de la norma ISAAR. Aquesta norma descriu "<i>corporate bodies, persons and families</i>".</p> <p>EAC té per objectiu la descripció d'entitats i persones, i el seu objecte equival doncs a aquesta entitat.</p> <p>El model MoReq recull el concepte d'agent únicament en funció de la seva relació amb un sistema d'informació, i ho fa a través de l'entitat E14.2.16 User.</p>

	<p>El model AGRkMS inclou l'entitat Agent, que defineix com “<i>A corporate entity, organisational element or system, or individual responsible for the performance of some business activity, including actions on records</i>”.</p> <p>El model e-EMGDE inclou l'entitat Agente, que pot ser una institució, un òrgan (administratiu), una persona o un dispositiu.</p> <p>La norma ISO 23081 contempla People (agents) com una de les cinc entitats principals del seu model.</p> <p>El model NEDA contempla l'entitat Agente, i la subdivideix en tres subtipus: Institución, Persona i Familia.</p> <p>El concepte d'agent a RiC és desenvolupat per vàries entitats. Aquest model no pretén desenvolupar cap tipologia d'agents, i tracta l'agent com una única entitat assimilable de forma general a persona jurídica. Malgrat RiC contempla també el concepte de “<i>mechanism</i>”, en aquest model s'infereix que cap mecanisme pot actuar sense programació/activació humana, i que sempre representa a un agent humà (l'agent competent en el procediment), per la qual cosa no s'aplica la identificació de mecanisme amb agent que proposa RiC.</p> <p>PROV contempla l'entitat prov:Agent, que defineix com “<i>something that bears some form of responsibility for an activity taking place, for the existence of an entity, or for another agent's activity</i>”. PROV descomposa aquesta entitat en tres subclasses: prov:Person, prov:Organization, i prov:SoftwareAgent. Aquesta darrera, que PROV defineix com “<i>running software</i>”, és equiparable al concepte de mecanisme de RiC.</p>
--	--

4.6.1.1.2 Atribut: Identificador

Identificador	ENT001-001
Nom formal	Identificador
Nom curt	Id
Domini	Identificador únic assignat a cada agent.
Tipus de dades	Alfanumèric
Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	<p>Aquest atribut es pot considerar equivalent a l'element 5.1.6 Identifiers for corporate bodies de la norma ISAAR.</p> <p>Correspon a l'element EAC <recordId>.</p> <p>El model MoReq recull aquest atribut en la metadada M14.4.100 System Identifier.</p>

	<p>Aquest atribut equival a la propietat 2 Identifier del model AGRkMS.</p> <p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE2 – Identificador.</p> <p>Equival a l'atribut identifier que preveu la norma ISO 23081.</p> <p>Correspon a l'atribut <AGE_ID> del model NEDA.</p> <p>Correspon a l'atribut genèric RiC-A22 Identifier.</p>
Exemples	20456

4.6.1.1.3 Atribut: Nom

Identificador	ENT001-002
Nom formal	Nom
Nom curt	Nom
Domini	Denominació formal de l'agent.
Tipus de dades	Alfanumèric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	<p>Aquest atribut es pot considerar equivalent a l'element 5.1.2 Authorized form(s) of name de la norma ISAAR. ISAAR ofereix altres opcions per a recollir noms alternatius, que recullen els elements 5.1.3 Parallel forms of name de la norma ISAAR, 5.1.4 Standardized forms of name according to other rules i 5.1.5 Other forms of name de la norma ISAAR.</p> <p>Aquest atribut equival a la propietat 3 Name del model AGRkMS.</p> <p>EAC disposa de l'element <nameEntry>, com a part de l'element <identity>, per descriure el nom d'un agent.</p> <p>El model MoReq recull aquest atribut en la metadada M14.4.104 Title.</p> <p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE3 – Nombre.</p> <p>Equival a l'atribut name que preveu la norma ISO 23081.</p> <p>Correspon a l'atribut <AGE_NOMBRE> del model NEDA.</p> <p>Correspon a l'atribut genèric RiC-A28 Name.</p>
Exemples	Servei de Personal

4.6.1.1.4 Atribut: Tipus

Identificador	ENT001-003
Nom formal	Tipus
Nom curt	Tipus
Domini	Tipus d'agent. La tipologia d'agents es defineix en funció del sistema de governança que es vulgui establir, i de l'abast d'aquest sistema de governança. Habitualment el tipus indica, en un primer nivell, si es tracta d'un agent propi de la persona jurídica de rang màxim en l'abast del sistema de governança, o bé d'un agent extern que és necessari referenciar per tal de donar coherència a tota la informació del sistema de governança. En un segon nivell, el tipus pot incloure el rang jeràrquic que l'agent ocupa dins la persona jurídica objecte de l'abast del sistema.
Tipus de dades	Alfanumèric categòric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	<p>L'atribut pot prendre els següents valors dins una jerarquia, que es pot anar desplegant en funció de la naturalesa, l'estructura i les relacions de cada ens:</p> <ul style="list-style-type: none"> • Entitat <ul style="list-style-type: none"> ○ Direcció ○ Servei ○ [Etc.] • Extern
Valor per defecte	No aplica.
Fonts	<p>Aquest atribut es pot considerar equivalent a l'element 5.1.1 Type of entity de la norma ISAAR. I també pot contenir informació relacionada amb l'element 5.2.4 Legal status.</p> <p>Aquest atribut equival a la propietat 14 Position del model AGRkMS.</p> <p>Es pot inferir una certa relació entre aquest atribut i l'element <legalStatus> del model EAC.</p> <p>Equival a l'atribut entity type que preveu la norma ISO 23081.</p> <p>Les relacions amb RiC són nombroses. La identificació més directa es pot fer amb l'atribut RiC-A12 Corporate body type, doncs és el que correspon a una categorització. Això no obstant, en el cas de determinades persones jurídiques també inclouria significació de RiC-A26 Legal status, el qual és aplicable entre d'altres a l'entitat RiC-E07 Agent. I en determinats casos també seria equiparable l'atribut RiC-A20 Family type.</p> <p>Quan representa rols de persones físiques, es podria equiparar també a RiC-A30 Occupation type. Però en aquest punt cal tenir present que amb l'atribut RiC-A15 Demographic Group el model RiC</p>

	permet descriure qualsevol grup humà amb característiques comunes, incloent també la possibilitat de definir perfils de persona o categories similars al rol. RiC el planteja com un atribut amb valors categòrics o controlats. En el cas d'aquest model, aquest atribut equival a les categories que es poden establir per als agents externs a l'organització.
Exemples	Servei

4.6.1.1.5 Atribut: Contacte

Identificador	ENT001-004
Nom formal	Dades de contacte
Nom curt	Contacte
Domini	Dades per contactar amb un representant humà de l'agent. Pot tractar-se d'un correu electrònic o qualsevol altre tipus d'adreça que faci possible una comunicació amb l'agent. Aquest atribut s'ha inclòs en el model com a resultat de necessitats específiques de gestió de la informació del Consorci d'Educació de Barcelona. No es considera determinant als efectes d'anàlisi de les dades recollides.
Tipus de dades	Usuari
Mín. ocurrencies	0
Màx. ocurrencies	5
Regles de validació	Es tracta d'un atribut complex de tipus persona física, que en funció del sistema on s'implementi pot incloure nom d'usuari, correu electrònic, telèfon de contacte, dades personals succintes, etc.
Valor per defecte	No aplica.
Fonts	EAC disposa de l'element <contactLine> per a descriure aquest atribut. Aquest atribut equival a la propietat 13 Contact del model AGrkMS . El model e-EMGDE contempla per aquest atribut la metadada eEMGDE10 - Contacto . Aquest atribut és equivalent a la metadada Point of Contact del model ONDC .
Exemples	Jordiserra18@outlook.es

4.6.1.1.6 Atribut: Observacions

Identificador	ENT001-005
Nom formal	Observacions
Nom curt	Observacions
Domini	Comentaris per definir l'abast i naturalesa de l'agent.

Tipus de dades	Alfanumèric descriptiu
Mín. ocurrències	0
Màx. ocurrències	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	<p>Aquest atribut pot contenir informació equivalent a l'element 5.2.2 History de la norma ISAAR. També en pot contenir l'element 5.2.8 General context.</p> <p>Aquest atribut equival a la propietat 5 Description del model AGRkMS.</p> <p>EAC disposa de l'element <description>, que inclou sub-elements que permeten descriure amb detall la història i característiques d'un agent, especialment els elements <abstract> i <biogHist>.</p> <p>El model MoReq recull aquest atribut en la metadada M14.4.16 Description.</p> <p>Equival a l'atribut description que preveu la norma ISO 23081.</p> <p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE5 - Descripción.</p> <p>Correspon a l'atribut <AGE_SINTESIS> del model NEDA.</p> <p>RiC recull el concepte de descripció o comentaris de caràcter genèric en l'atribut RiC-A16 Descriptive note. També pot incloure puntualment informacions descriptives de la història de l'agent que equivalen a RiC-A21 History.</p>
Exemples	Servei que s'ocupa de la gestió dels recursos humans dins el Consorci d'Educació de Barcelona i en els centres educatius de titularitat pública.

4.6.1.1.7 Atribut: Data creació

Identificador	ENT001-006
Nom formal	Data de creació
Nom curt	Data creació
Domini	Data oficial acreditable de creació o constitució de l'agent (en el cas d'una persona jurídica).
Tipus de dades	Data
Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	Data vàlida.
Valor per defecte	No aplica.
Fonts	Aquest atribut es pot considerar equivalent a l'element 5.2.1 Dates of existence de la norma ISAAR .

	<p>Correspon a l'element <date> del model EAC.</p> <p>El model MoReq recull aquest atribut en la metadada M14.4.9 Created timestamp, tot i que també s'hi pot aplicar la metadada M14.4.61 originated Date/time.</p> <p>Aquest atribut equival a la propietat 4 Date range del model AGRkMS.</p> <p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE4.1 - Fecha inicio.</p> <p>Correspon a l'atribut <AGE_FECHA> del model NEDA.</p> <p>Degut a la complexitat inherent a la descripció cronològica, RiC ha optat per tractar les dates com a entitat enlloc d'atributs. Amb aquest objectiu disposa d'una entitat, RiC-E18 Date, que es descompon en tres sub-entitats: RiC-E19 Single date, RiC-E20 Date range, i RiC-E21 Date set. En aquest model, tots els conjunts de data inici + data fi els considerarem equivalents a l'atribut RiC-A29 Normalized date dins l'entitat RiC-E20 Date range.</p>
Exemples	12 de gener de 2008.

4.6.1.1.8 Atribut: Data supressió

Identificador	ENT001-007
Nom formal	Data de supressió
Nom curt	Data supressió
Domini	<p>Data en la qual l'agent cessa oficialment en la seva activitat, es dissol o desapareix. Mentre no s'informa aquesta data es considera que l'agent està vigent.</p> <p>Les discontinuïtats en la vida o existència de l'agent es faran constar en l'atribut observacions.</p>
Tipus de dades	Data
Mín. ocurrencies	0
Màx. ocurrencies	1
Regles de validació	Data vàlida.
Valor per defecte	No aplica.
Fonts	<p>Aquest atribut es pot considerar equivalent a l'element 5.2.1 Dates of existence de la norma ISAAR.</p> <p>Correspon a l'element <date> del model EAC.</p> <p>El model MoReq recull aquest atribut en la metadada M14.4.5 Closed timestamp.</p> <p>Aquest atribut equival a la propietat 4 Date range del model AGRkMS.</p>

	<p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE4.2 - Fecha fin.</p> <p>Correspon a l'atribut <AGE_FECHA> del model NEDA.</p> <p>Equival a l'atribut RiC-A29 Normalized date dins l'entitat RiC-E20 Date range. Vegeu nota explicativa a l'apartat Fonts de l'atribut ENT001-006.</p>
Exemples	31 de desembre de 2021.

4.6.1.2 Entitat: Procés

4.6.1.2.1 Descripció de l'entitat

Identificador	ENT002
Nom formal	Procés
Nom curt	Procés
Domini	Definirem procés com el conjunt d'activitats d'un o diversos agents adreçades a aconseguir un objectiu comú, generades per una entrada (<i>input</i>) d'un proveïdor, i orientades a la producció d'un producte o servei (<i>output</i>) cap a un client (intern o extern).
Atributs	ENT002-001 Identificador ENT002-002 Nom ENT002-003 Descripció ENT002-004 Tipus ENT002-005 Família ENT002-006 Àmbit ENT002-007 Tipus iniciació ENT002-008 Tipus finalització ENT002-009 Tipus recurs ENT002-010 Requisits a acreditar ENT002-011 Terminis rellevants ENT002-012 Necessitat de pagament ENT002-013 Data creació ENT002-014 Data supressió
Relacions	REL002 Té com responsable un agent REL002 Pot tenir com a destinatari un o més agents REL008 Pot estar relacionat amb un o més processos REL009 Es descompon en una o més tasques REL010 Pot ser regulat per una o més disposicions normatives REL013 Dona origen a una o més sèries documentals REL016 Genera una o més evidències
Fonts	<p>La norma ISDF té per objectiu la descripció de funcions desenvolupades en el si de les organitzacions. ISDF utilitza un concepte ampli de funció, que inclou "<i>not only function but also any of the subdivisions of a function such as subfunction, business process, activity, task, transaction or other term in international, national or local usage</i>". Per tant, la coincidència amb l'entitat procés és exacta.</p> <p>Però a més de la norma ISDF, trobem també a la norma ISAAR un element que permet definir funcions i activitats. Aquest element és 5.2.5 Functions, occupations and activities.</p> <p>MoReq no recull específicament cap entitat corresponent a funció, activitat o procés, però inclou dins la definició de l'entitat E14.2.2 Class la seva correspondència amb les classificacions de negoci que es poden aplicar a agrupacions documentals o a documents en</p>

	<p>concret, i que en el redactat de MoReq es defineixen com les funcions, activitats i processos de negoci de l'organització.</p> <p>El model AGRkMS contempla l'entitat Business, que defineix com “<i>A business function, activity or transaction performed by, or assigned to, an organisation or its employees</i>”.</p> <p>La norma ISO 23081 contempla Business com una de les cinc entitats principals del seu model, dins de la qual inclou els processos de gestió documental (Records management business).</p> <p>El model e-EMGDE contempla també una entitat que anomena Actividad, la qual descompon en quatre nivells de concreció: Función marco, Función, Actividad i Acción.</p> <p>El model NEDA contempla l'entitat procés dins de la jerarquia habitual. Ho fa amb l'entitat Función y sus divisiones, la qual descompon en quatre subtipus: Función, División de función, Actividad o proceso i Acción.</p> <p>L'entitat del model RiC que més s'apropa al concepte de procés és RiC-E15 Activity, una subclasse de RiC-E14 Event, que equival en la seva descripció al concepte de funció i/o activitat propi dels quadres de classificació funcional arxivístics.</p> <p>DCAT utilitza un element de l'ontologia PROV que és prov:wasGeneratedBy, i que correspon a l'activitat que ha generat el <i>dataset</i>, o bé que li proveeix el context de negoci. És per tant un element clarament assimilable a procés.</p> <p>Una de les tres entitats principals del model PROV és prov:Activity, que correspon clarament amb el concepte de procés, i que PROV defineix com “<i>something that occurs over a period of time and acts upon or with entities; it may include consuming, processing, transforming, modifying, relocating, using, or generating entities</i>”. Aquestes entitats a les quals fa referència PROV poden ser equivalents als objectes descrits del nostre model (ENT014), que són també objecte de gestió o operació per part dels processos identificats. Però hi ha una segona possible lectura que es detalla en l'apartat Fonts de l'entitat ENT014.</p>
--	--

4.6.1.2.2 Atribut: Identificador

Identificador	ENT002-001
Nom formal	Identificador
Nom curt	Id
Domini	Codi que identifica unívocament el procés dins el conjunt de processos objecte de governança integrada.

	En termes arxivístics, el codi de procés també rep el nom de “codi de classificació”, i és sovint recollit en un quadre de classificació funcional.
Tipus de dades	Alfanumèric
Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	Es tracta d'un codi qualificat, compost de tres lletres, que identifiquen la família del procés, i un número seqüencial dins aquella família de processos.
Valor per defecte	No aplica.
Fonts	<p>Aquest atribut es pot considerar equivalent a l'element 5.1.5 Classification de la norma ISDF, tot i que sense vinculació amb cap sistema de classificació arxivístic, a diferència de l'esperit de la ISDF. I també, de forma més precisa, a l'element 5.4.1 Function description identifier de la norma ISDF.</p> <p>El model MoReq recull aquest atribut en la metadada M14.4.100 System Identifier.</p> <p>Aquest atribut equival a la propietat 2 Identifier del model AGRkMS.</p> <p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE2 – Identificador.</p> <p>Equival a l'atribut identifíer que preveu la norma ISO 23081.</p> <p>Correspon a l'atribut <FUNC_ID> del model NEDA.</p> <p>Correspon a l'atribut genèric RiC-A22 Identifier.</p>
Exemples	ORG003

4.6.1.2.3 Atribut: Nom

Identificador	ENT002-002
Nom formal	Nom
Nom curt	Nom
Domini	Nom formal amb el qual es descriu el procés.
Tipus de dades	Alfanumèric
Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	<p>Aquest atribut es pot considerar equivalent a l'element 5.1.2 Authorised form(s) of name de la norma ISDF. Si hi hagués més d'una denominació, ISDF permet utilitzar també els elements 5.1.3 Parallel form(s) of name i 5.1.4 Other form(s) of name.</p> <p>El model MoReq recull aquest atribut en la metadada M14.4.104 Title.</p>

	<p>Aquest atribut equival a la propietat 3 Name del model AGRkMS.</p> <p>Equival a l'atribut name que preveu la norma ISO 23081.</p> <p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE3 – Nombre.</p> <p>Correspon a l'atribut <FUNC_NOMBRE> del model NEDA.</p> <p>Correspon a l'atribut genèric RiC-A28 Name.</p>
Exemples	Selecció de personal

4.6.1.2.4 Atribut: Descripció

Identificador	ENT002-003
Nom formal	Descripció
Nom curt	Descripció
Domini	Descripció de l'objectiu, desenvolupament bàsic i altres característiques del procés.
Tipus de dades	Alfanumèric descriptiu
Mín. ocurrències	0
Màx. ocurrències	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	<p>Aquest atribut es pot considerar equivalent a l'element 5.2.2 Description de la norma ISDF.</p> <p>El model MoReq recull aquest atribut en la metadada M14.4.16 Description.</p> <p>Aquest atribut equival a la propietat 5 Description del model AGRkMS.</p> <p>Equival a l'atribut description que preveu la norma ISO 23081.</p> <p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE5 - Descripción.</p> <p>Correspon a l'atribut <FUNC_SINTESIS> del model NEDA.</p> <p>Equival també a RiC-A16 Descriptive note.</p>
Exemples	Consisteix en l'anàlisi i la millora de processos, a partir d'un encàrrec fet per un òrgan directiu o de la petició directa per part del responsable del procés afectat. Es pot portar a terme amb diferents metodologies, i finalitza amb un informe de millores del procés i, si escau, amb la implantació efectiva d'aquestes millores.

4.6.1.2.5 Atribut: Tipus

Identificador	ENT002-004
Nom formal	Tipus de procés
Nom curt	Tipus
Domini	Assignació de tipus al procés d'acord amb la tipologia que diferencia entre processos estratègics, clau i de suport.
Tipus de dades	Alfanumèric categòric
Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	Ha de contenir un dels següents valors: <ul style="list-style-type: none">• Estratègic• Clau• De suport
Valor per defecte	No aplica.
Fonts	Correspon a l'atribut <FUNC_TIPOENT> del model NEDA.
Exemples	De suport

4.6.1.2.6 Atribut: Família

Identificador	ENT002-005
Nom formal	Família del procés
Nom curt	Família
Domini	Descriptor que classifica el procés en funció de la seva similitud amb d'altres processos del mateix àmbit funcional, principalment pel seu objecte i/o finalitat.
Tipus de dades	Alfanumèric categòric
Mín. ocurrències	0
Màx. ocurrències	1
Regles de validació	Ha de contenir valors de la llista de famílies que es va elaborant a mesura que s'identifiquen i caracteritzen els processos.
Valor per defecte	No aplica.
Fonts	RiC contempla la possibilitat de desplegar les activitats de forma taxonòmica, i recollir aquesta categorització en l'atribut RiC-A02 Activity type . En aquest model es treballa amb un sol nivell d'agregació, el de família de procés, que correspon a aquest atribut de RiC .
Exemples	Organitzatius

4.6.1.2.7 Atribut: Àmbit

Identificador	ENT002-006
Nom formal	Àmbit del procés
Nom curt	Àmbit

Domini	Àmbit orgànic en el qual s'inclou el procés en funció de la naturalesa del seu responsable. Aquest atribut és especialment important en aquells casos en què l'abast del sistema de governança és complex i inclou diversos ens o agents.
Tipus de dades	Alfanumèric categòric
Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	Ha de contenir un valor de l'entitat "Agents", seleccionat entre els agents de rang superior (no dependents de cap altre agent).
Valor per defecte	No aplica.
Fonts	El model e-EMGDE contempla per aquest atribut la metadada eEMGDE7 - Àmbito .
Exemples	ConSORCI d'Educació de Barcelona

4.6.1.2.8 Atribut: Tipus iniciació

Identificador	ENT002-007
Nom formal	Tipus d'iniciació
Nom curt	Iniciació
Domini	Tipus de fet que actua com a disparador o <i>trigger</i> del procés, i que n'origina o activa una instància concreta. Aquest atribut s'ha inclòs en el model com a resultat de necessitats específiques de gestió de la informació del Consorci d'Educació de Barcelona. No es considera determinant als efectes d'anàlisi de les dades recollides.
Tipus de dades	Alfanumèric categòric
Mín. ocurrències	0
Màx. ocurrències	1
Regles de validació	Ha de contenir un dels següents valors: <ul style="list-style-type: none"> • D'ofici • A instància de part • Per denúncia
Valor per defecte	No aplica.
Fonts	Es pot considerar equivalent a l'atribut trigger que inclou la norma ISO 23081 .
Exemples	D'ofici

4.6.1.2.9 Atribut: Tipus finalització

Identificador	ENT002-008
Nom formal	Tipus de finalització
Nom curt	Finalització

Domini	Tipus de fet que posa fi al procés, i en defineix la seva darrera tasca. Pot tractar-se de la disponibilitat d'un producte, de la finalització de la prestació d'un servei, o de la finalització d'una acció de control. Aquest atribut s'ha inclòs en el model com a resultat de necessitats específiques de gestió de la informació del Consorci d'Educació de Barcelona. No es considera determinant als efectes d'anàlisi de les dades recollides.
Tipus de dades	Alfanumèric categòric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	Ha de contenir valors de la llista de finalitzacions que es va elaborant a mesura que s'identifiquen i caracteritzen els processos. Aquesta és una llista que es va elaborant dinàmicament a mesura que s'identifiquen nous tipus de finalització.
Valor per defecte	No aplica.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	Lliurament de l'informe de millores

4.6.1.2.10 Atribut: Tipus recurs

Identificador	ENT002-009
Nom formal	Tipus de recurs
Nom curt	Recurs
Domini	Tipus d'acció que pot revisar el resultat del procés o en pot reactivar l'execució per tal d'esmenar-ne la finalització. El concepte de recurs és propi del dret administratiu, però també és aplicable a qualsevol mena de revisió d'un procés posterior a la seva finalització i tancament formal. Aquest atribut s'ha inclòs en el model com a resultat de necessitats específiques de gestió de la informació del Consorci d'Educació de Barcelona. No es considera determinant als efectes d'anàlisi de les dades recollides.
Tipus de dades	Alfanumèric categòric
Mín. ocurrencies	0
Màx. ocurrencies	1
Regles de validació	En el cas del sector públic, el valor inclourà totes les modalitats de recurs o procediment de revisió que contempli la normativa aplicable. En el cas del sector privat o quan no sigui aplicable el procediment administratiu, el valor inclourà els supòsits de reclamació i reparació que es puguin donar en cada cas.
Valor per defecte	No aplica.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	Recurs d'alçada

4.6.1.2.11 Atribut: Requisits a acreditar

Identificador	ENT002-010
Nom formal	Requisits a acreditar
Nom curt	Requisits
Domini	<p>Requisits que cal que es compleixin per tal que el procés es pugui iniciar. L'acreditació de requisits és un concepte propi del dret administratiu, però que també és aplicable a qualsevol mena de control previ a la iniciació d'un procés, i que en pot determinar, en cas d'incompliment, la seva finalització anormal o anticipada.</p> <p>Aquest atribut s'ha inclòs en el model com a resultat de necessitats específiques de gestió de la informació del Consorci d'Educació de Barcelona. No es considera determinant als efectes d'anàlisi de les dades recollides.</p>
Tipus de dades	Alfanumèric categòric
Mín. ocurrencies	0
Màx. ocurrencies	1
Regles de validació	Ha de contenir valors de la llista de requisits que es va elaborant a mesura que s'identifiquen i caracteritzen els processos.
Valor per defecte	No aplica.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	Família monoparental

4.6.1.2.12 Atribut: Terminis rellevants

Identificador	ENT002-011
Nom formal	Terminis rellevants
Nom curt	Terminis
Domini	<p>Fites temporals que ha de complir un procés durant la seva execució. Poden aplicar al desenvolupament del procés, a determinades tasques del procés, o al temps estipulat per a la seva finalització.</p> <p>Aquest atribut s'ha inclòs en el model com a resultat de necessitats específiques de gestió de la informació del Consorci d'Educació de Barcelona. No es considera determinant als efectes d'anàlisi de les dades recollides.</p>
Tipus de dades	Alfanumèric
Mín. ocurrencies	0
Màx. ocurrencies	1
Regles de validació	Ha de contenir valors fixats sobre calendari. Poden ser dates concretes o terminis a calcular a partir d'esdeveniments.
Valor per defecte	No aplica.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	Primer dilluns de juny

4.6.1.2.13 Atribut: Necessitat de pagament

Identificador	ENT002-012
Nom formal	Necessitat de pagament
Nom curt	Pagament
Domini	<p>Indicació de si un procés requereix un pagament com a condició per a la seva iniciació, desenvolupament o finalització.</p> <p>Aquest atribut s'ha inclòs en el model com a resultat de necessitats específiques de gestió de la informació del Consorci d'Educació de Barcelona. No es considera determinant als efectes d'anàlisi de les dades recollides.</p>
Tipus de dades	Booleà
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	Només pot prendre dos valors: sí o no.
Valor per defecte	"No"
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	No

4.6.1.2.14 Atribut: Data creació

Identificador	ENT002-013
Nom formal	Data de creació
Nom curt	Creació
Domini	<p>Data en la qual un procés és reconegut formalment com a tal dins el mapa de processos de l'organització. Aquesta data pot provenir de tres fonts:</p> <ul style="list-style-type: none">• D'una norma jurídica (en el cas de processos que requereixen una aprovació de tipus administratiu o jurídic, o que estan reflectits en l'atribució formal de competències a un determinat agent).• D'un manual de procediment (en el cas que existeixi dins l'organització un procediment formal d'aprovació dels processos).• D'un mapa o registre de processos (en el cas que l'organització disposi, en el seu sistema de governança, d'un instrument d'aquest tipus). <p>En cas que la data no es pugui determinar per la inexistència de les fonts mencionades, la data que es consignarà com a data d'inici del procés serà la data de la creació de l'agent responsable d'aquell procés.</p>
Tipus de dades	Data
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	Data vàlida.

Valor per defecte	No aplica.
Fonts	<p>Aquest atribut es pot considerar equivalent a l'element 5.2.1 Date de la norma ISDF.</p> <p>El model MoReq recull aquest atribut en la metadada M14.4.9 Created timestamp, tot i que també s'hi pot aplicar la metadada M14.4.61 originated Date/time.</p> <p>Aquest atribut equival a la propietat 4 Date range del model AGRkMS.</p> <p>Equival a l'atribut date/time que preveu la norma ISO 23081.</p> <p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE4.1 - Fecha inicio.</p> <p>Correspon a l'atribut <FUNC_FECHA> del model NEDA.</p> <p>Equival a l'atribut RiC-A29 Normalized date dins l'entitat RiC-E20 Date range. Vegeu nota explicativa a l'apartat Fonts de l'atribut ENT001-006.</p>
Exemples	12 de gener de 2008.

4.6.1.2.15 Atribut: Data supressió

Identificador	ENT002-014
Nom formal	Data de supressió
Nom curt	Supressió
Domini	<p>Data en la qual el procés deixa de portar-se a terme, o és substituït per una nova lectura funcional de l'organització. Aquesta data, que habitualment no té una fixació formal, serà detectada a partir de revisions periòdiques del mapa de processos i de l'anàlisi de possibles substitucions d'un procés per un altre, arran de l'aparició de nous processos, i serà certificada mitjançant treball de camp.</p> <p>En el cas de la desaparició de l'agent responsable d'un procés, s'avaluarà la continuïtat d'aquell procés sota la responsabilitat d'un nou agent, i en cas que no es pugui determinar aquesta continuïtat es donarà per suprimit el procés en la mateixa data en què hagi desaparegut l'agent responsable.</p>
Tipus de dades	Data
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	Data vàlida.
Valor per defecte	No aplica.
Fonts	Aquest atribut es pot considerar equivalent a l'element 5.2.1 Date de la norma ISDF .

	<p>El model MoReq recull aquest atribut en la metadada M14.4.5 Closed timestamp.</p> <p>Aquest atribut equival a la propietat 4 Date range del model AGRkMS.</p> <p>Equival a l'atribut date/time que preveu la norma ISO 23081.</p> <p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE4.2 - Fecha fin.</p> <p>Correspon a l'atribut <FUNC_FECHA> del model NEDA.</p> <p>Equival a l'atribut RiC-A29 Normalized date dins l'entitat RiC-E20 Date range. Vegeu nota explicativa a l'apartat Fonts de l'atribut ENT001-006.</p>
Exemples	31 de desembre de 2021.

4.6.1.3 Entitat: Tasca

4.6.1.3.1 Descripció de l'entitat

Identificador	ENT003
Nom formal	Tasca
Nom curt	Tasca
Domini	Tasca és cada acció (o grup d'accions) realitzada per un agent dins un procés. Totes les tasques d'un procés estan relacionades per una relació de causa-efecte, relacions que poden implicar bifurcacions en funció de determinades condicions.
Atributs	ENT003-001 Identificador ENT003-002 Nom ENT003-003 Descripció ENT003-004 Tipus executiu ENT003-005 Tipus en el flux ENT003-006 Condició de bifurcació ENT003-007 Data creació ENT003-008 Data supressió
Relacions	REL003 És executada per un agent REL009 Forma part d'un procés REL011 Pot ser antecedent de una o més tasques REL011 Pot ser conseqüent a una o més tasques REL017 Pot generar/rebre una o més evidències
Fonts	<p>La norma ISDF, malgrat que el seu objectiu és definir funcions i activitats a un nivell mínimament alt (equivalent a procés), també inclou la possibilitat de ser aplicada al nivell de tasca ("<i>task</i>" o "<i>transaction</i>"), com detalla la pròpia norma en la definició de l'abast.</p> <p>MoReq disposa de l'element E14.2.8 Event, que defineix com "<i>the outcome of a function that was performed previously and is retained to show the history of an entity</i>". Aquest element s'orienta tanmateix a la descripció d'esdeveniments reals ja passats als efectes de pista d'auditoria, i no a la tipificació d'esdeveniments dins un procés. La correspondència no és doncs directa.</p> <p>El model NEDA contempla l'entitat tasca dins de la jerarquia habitual pròpia de la ISO 15489, com a subtipus de l'entitat Función y sus divisiones, un dels subtipus de la qual és Acción.</p> <p>Les mencions a "activitat" que trobem en la resta de models analitzats fan referència a nivells superiors, i no els considerem equivalents al nivell de "tasca" del nostre model.</p>

4.6.1.3.2 Atribut: Identificador

Identificador	ENT003-001
Nom formal	Identificador
Nom curt	Id
Domini	Identificador únic de la tasca dins el procés del qual forma part.
Tipus de dades	Alfanumèric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	L'identificador de cada tasca s'assigna com un codi qualificat, compostat del codi de procés (6 dígit) seguit, després d'un guió, per un número seqüencial de tres dígit.
Valor per defecte	No aplica.
Fonts	Aquest atribut es pot considerar equivalent a l'element 5.4.1 Function description identifier de la norma ISDF . El model MoReq recull aquest atribut en la metadada M14.4.100 System Identifier . Correspon a l'atribut < FUNC_ID > del model NEDA . Correspon a l'atribut genèric RiC-A22 Identifier .
Exemples	GRH001-003

4.6.1.3.3 Atribut: Nom

Identificador	ENT003-002
Nom formal	Nom
Nom curt	Nom
Domini	Descripció succinta de la tasca. És el nom que descriu aquesta tasca en les guies de procediment o en les diagramacions del procés.
Tipus de dades	Alfanumèric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	Aquest atribut es pot considerar equivalent a l'element 5.1.2 Authorised form(s) of name de la norma ISDF . El model MoReq recull aquest atribut en la metadada M14.4.104 Title . Correspon a l'atribut < FUNC_NOMBRE > del model NEDA . Correspon a l'atribut genèric RiC-A28 Name .
Exemples	Elaborar informe

4.6.1.3.4 Atribut: Descripció

Identificador	ENT003-003
Nom formal	Descripció
Nom curt	Descripció
Domini	Descripció completa de l'objecte, desenvolupament i forma d'execució de la tasca. Pot incloure informació sobre les excepcions, condicions o terminis d'execució de la tasca.
Tipus de dades	Alfanumèric descriptiu
Mín. ocurrences	0
Màx. ocurrences	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	<p>Aquest atribut es pot considerar equivalent a l'element 5.2.2 Description de la norma ISDF.</p> <p>MoReq contempla una metadada M14.4.25 Event Comment per afegir comentaris a un esdeveniment, tot i que la correspondència conceptual amb aquest atribut no és exacta (són comentaris sobre l'execució, no de definició de la tasca).</p> <p>Correspon a l'atribut <FUNC_SINTESIS> del model NEDA.</p> <p>Equivalent a RiC-A16 Descriptive note.</p>
Exemples	Elaborar un informe textual que descrigui la idoneïtat del candidat per a la plaça o lloc de treball a ocupar.

4.6.1.3.5 Atribut: Tipus executiu

Identificador	ENT003-004
Nom formal	Tipus executiu
Nom curt	Tipus
Domini	<p>Assignació d'un tipus de tasca en funció de l'autonomia en la seva execució, valorant si pot ser una tasca que es realitzi de forma desatesa (automàtica), o requereix de l'acció voluntària d'un agent humà.</p> <p>Aquest tipus només s'assigna a les tasques que descriuen accions.</p>
Tipus de dades	Alfanumèric categòric
Mín. ocurrences	0
Màx. ocurrences	1
Regles de validació	<p>L'atribut ha de prendre un dels següents valors:</p> <ul style="list-style-type: none">• Automàtica• Manual• Interacció home-màquina
Valor per defecte	No aplica.

Fonts	Considerem una possible equivalència amb l'atribut RiC-A18 Event type . Correspon a l'atribut <FUNC_TIPOENT> del model NEDA .
Exemples	Manual

4.6.1.3.6 Atribut: Tipus en el flux

Identificador	ENT003-005
Nom formal	Tipus en el flux
Nom curt	Tipus flux
Domini	Assignació d'un tipus de tasca als objectes de completar la lògica d'execució del flux i fer-ne possible una diagramació automàtica. Això inclou les tasques que descriuen accions, i també les tasques que descriuen condicions, bifurcacions o terminis (falses tasques). Aquest atribut s'ha inclòs en el model per tal de facilitar la generació automatitzada de les diagramacions dels processos, funcionalitat de l'eina CATPROC explicada en l'apartat 4.3.3 d'aquest capítol. No es considera determinant als efectes d'anàlisi de les dades recollides.
Tipus de dades	Alfanumèric categòric
Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	L'atribut ha de prendre un dels següents valors: ²¹⁴ <ul style="list-style-type: none"> • Inici del procés • Tasca • Subprocés (vinclat amb) • <i>Gateway</i> • <i>Gateway</i> exclusiu • <i>Gateway</i> inclusiu • <i>Gateway</i> paral·lel • Esdeveniment • Termini • Missatge • Evidència d'entrada • Evidència de sortida • Interrupció del procés • Fi del procés
Valor per defecte	"Tasca"
Fonts	Considerem una possible equivalència amb l'atribut RiC-A18 Event type .
Exemples	Termini

²¹⁴ Aquests atributs són propis de la notació BPMN (*Business Process Model and Notation*), estàndard al qual hem fet referència en nota anterior, mantingut per l'*Object Management Group* (OMG), i actualment en la versió 2.0.2.

4.6.1.3.7 Atribut: *Condició de bifurcació*

Identificador	ENT003-006
Nom formal	Condició de bifurcació
Nom curt	Condició
Domini	<p>Literal que descriu el valor de condició que determina l'execució d'una tasca quan l'antecedent és una bifurcació (<i>gateway</i>).</p> <p>Com en el cas anterior, aquest atribut s'ha inclòs en el model per tal de facilitar la generació automatitzada de les diagramacions dels processos, funcionalitat de l'eina CATPROC explicada en l'apartat 4.3.3 d'aquest capítol. No es considera determinant als efectes d'anàlisi de les dades recollides.</p>
Tipus de dades	Alfanumèric
Mín. ocurrencies	0
Màx. ocurrencies	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	
Exemples	Si el termini és superior a 5 dies

4.6.1.3.8 Atribut: *Data creació*

Identificador	ENT003-007
Nom formal	Data de creació
Nom curt	Data creació
Domini	La data de creació d'una tasca coincidirà amb la data de creació del procés quan es tracti del disseny inicial, o amb la data de la modificació, redisseny o nou versionat del procés.
Tipus de dades	Data
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	Data vàlida.
Valor per defecte	No aplica.
Fonts	<p>Aquest atribut es pot considerar equivalent a l'element 5.2.1 Date de la norma ISDF.</p> <p>Correspon a l'atribut <FUNC_FECHA> del model NEDA.</p> <p>Equival a l'atribut RiC-A29 Normalized date dins l'entitat RiC-E20 Date range. Vegeu nota explicativa a l'apartat Fonts de l'atribut ENT001-006.</p>
Exemples	12 de gener de 2008.

4.6.1.3.9 Atribut: Data supressió

Identificador	ENT003-008
Nom formal	Data de supressió
Nom curt	Data supressió
Domini	La data de supressió serà la data de modificació, redisseny o nou versionat del procés que anul·li o modifiqui substancialment aquella tasca, o bé la data de supressió de tot el procés.
Tipus de dades	Data
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	Data vàlida.
Valor per defecte	No aplica.
Fonts	<p>Aquest atribut es pot considerar equivalent a l'element 5.2.1 Date de la norma ISDF.</p> <p>Correspon a l'atribut <FUNC_FECHA> del model NEDA.</p> <p>Equival a l'atribut RiC-A29 Normalized date dins l'entitat RiC-E20 Date range. Vegeu nota explicativa a l'apartat Fonts de l'atribut ENT001-006.</p>
Exemples	31 de desembre de 2021.

4.6.1.4 Entitat: Evidència

4.6.1.4.1 Descripció de l'entitat

Identificador	ENT004
Nom formal	Evidència
Nom curt	Evidència
Domini	<p>En el marc d'aquest model, definim evidència com qualsevol agregat de dades amb significació pròpia/individual que pot aparèixer en el curs d'un procés com a resultat de l'execució d'una tasca.</p> <p>Aquest agregat pot ser permanent o temporal (circumstancial). Si és permanent, pren la forma habitualment d'un document-objecte (en suport físic o en fitxer informàtic) que, amb posterioritat al seu ús primer, resta emmagatzemat en algun repositori. Si és temporal, pren la forma habitualment d'una pantalla d'entrada i/o sortida de dades, inclosa en una aplicació informàtica (o sistema d'informació), que mostra dades que estan emmagatzemades en base de dades.</p> <p>L'origen de l'evidència pot estar en el propi procés o en un procés anterior.</p>
Atributs	<p>ENT004-001 Identificador</p> <p>ENT004-002 Nom</p> <p>ENT004-003 Tipus</p> <p>ENT004-004 Caràcter</p> <p>ENT004-005 Format</p> <p>ENT004-006 Autenticació</p> <p>ENT004-007 Registre d'entrada</p> <p>ENT004-008 Registre de sortida</p> <p>ENT004-009 Essencial</p> <p>ENT004-010 Data creació</p> <p>ENT004-011 Data supressió</p>
Relacions	<p>REL007 És creada/rebuda per un agent</p> <p>REL016 És generada per un procés</p> <p>REL017 És generada/rebuda per una o més tasques</p> <p>REL018 Està continguda en una sèrie documental</p> <p>REL019 Està continguda en un sistema d'informació (tipus "dades")</p> <p>REL020 Conté una o més dades elementals</p> <p>REL021 Es pot formalitzar amb un o més models de document</p>
Fonts	<p>ISAD es planteja com un model de descripció multinivell, que permet descriure des del nivell de fons fins al nivell de document, que defineix com la unitat arxivística intel·lectualment indivisible (ISAD l'anomena ítem, que en la traducció espanyola s'ha anomenat unitat documental simple o document). L'entitat document es fa explícita a través de l'element 3.1.4 Level of description.</p> <p>EAD també contempla el document de forma explícita, citat com un dels nivells de descripció (atribut level) de l'element <archdesc>, en concret l'element ítem, malgrat que fent referència a documents</p>

	<p>específics, no a tipus de document de circuit, que és el que descriu aquesta entitat.</p> <p>El model AGRkMS inclou l'entitat record, que tant es pot assimilar a document com a expedient, tal com indica en la descripció: “[...] <i>A record may comprise an electronic or paper-based document or group of aggregated documents</i>”. La seva relació amb l'entitat Evidència no és per tant directa.</p> <p>El model e-EMGDE també inclou aquesta entitat, amb la denominació Documento, que descriu d'acord amb la definició habitual de document administratiu o d'arxiu: “<i>Información estructurada en cualquier formato creada, recibida y mantenida como evidencia por una organización o persona en cumplimiento de obligaciones legales o para actuaciones de gestión</i>”. De la mateixa manera que en el cas del AGRkMS contempla diferents possibles agrupacions d'aquesta entitat (que anomena categories), i que es corresponen amb els nivells arxivístics de grup de fons, fons, sèrie (documental), agregació i expedient, a part del nivell de document simple.</p> <p>La norma ISO 23081 contempla Records com una de les cinc entitats principals del seu model.</p> <p>MoReq recull aquest concepte en l'entitat E14.2.12 Record, als efectes de descripció de documents o evidències compostades per qualsevol element tecnològic.</p> <p>El model NEDA contempla l'entitat Documentos de archivo, la qual subdivideix en deu subtipus: Grupo de fondos, Fondo, División de fondo o de grupo de fondo, Serie, Subserie, Fracción de serie/subserie, Unidad documental, Colección, División de colección i Componente documental. En aquest cas l'entitat correspon al subtipus Unidad documental.</p> <p>Aquesta entitat correspon al concepte recollit en l'entitat RiC-E04 Record, o document de procés. No descriu per tant una entitat física, sinó lògica. L'entitat física és la que es recull en l'entitat RiC-E06 Instantiation. Però l'origen de la conceptualització de RiC i la que recollim en el nostre model és diferent, per la qual cosa no hem d'inferir una correspondència exacta entre la nostra entitat ENT004 Evidència i l'entitat RiC-E04 Record. Quan RiC descriu ambdues entitats es fa palès que s'inspira en el model PREMIS per a la preservació digital (<i>PREMIS Data Dictionary for Preservation Metadata</i>, 2015), considerant RiC-E04 Record l'equivalent del concepte PREMIS d'entitat intel·lectual, i RiC-E06 Instantiation l'equivalent del concepte PREMIS de representació.</p> <p>DCAT, malgrat que té per objecte descriure conjunts de dades preferentment estructurades, inclou conceptualment qualsevol forma documental que pugui esdevenir una font de dades. Aquesta</p>
--	--

	<p>visió és hereva del fet que un <i>data lake</i> pot contenir qualsevol mena d'objecte digital, contingui aquesta informació estructurada o no estructurada. Per aquest motiu, atribuïm metadades de DCAT tant a l'entitat sistema d'informació com a l'entitat evidència.</p> <p>Les evidències, siguin documents-objecte o estiguin basades en dades, son considerades actius d'informació ("<i>data assets</i>") en el marc de l'esquema ONDC, cosa la qual permet la seva descripció en catàlegs de dades.</p>
--	---

4.6.1.4.2 Atribut: Identificador

Identificador	ENT004-001
Nom formal	Identificador
Nom curt	ID
Domini	<p>Codi que identifica de forma unívoca l'evidència. La pauta d'assignació de codi inclou:</p> <ul style="list-style-type: none"> • El codi del procés en el curs del qual s'ha generat l'evidència • Un prefix format per la lletra "E" • Un número consecutiu que identifica l'evidència dins el seu procés, i que segueix l'ordre d'aparició de les evidències dins el procés • Quan un procés genera diverses còpies o exemplars d'un mateix document o pantalla de dades, s'utilitza una lletra, en minúscula, que identifica cada exemplar de l'evidència dins el procés.
Tipus de dades	Alfanumèric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	<p>Aquest atribut equival a la propietat 2 Identifier del model AGRkMS.</p> <p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE2 – Identificador.</p> <p>Equival a l'atribut identifier que preveu la norma ISO 23081.</p> <p>Correspon a l'atribut genèric RIC-A22 Identifier.</p> <p>Correspon a l'element ISAD 3.1.1 Reference code.</p> <p>EAD disposa de diferents elements per a descriure l'identificador d'una agrupació documental: <recordid> i <unitid>.</p> <p>Correspon a l'atribut <DOC_ID> del model NEDA.</p>

	<p>DCAT utilitza l'element de Dublin Core dct:identifïer per assignar un identificador ùnic a cada recurs d'un catàleg.</p> <p>Aquest atribut és equivalent a la metadada Identifïer del model ONDC.</p>
Exemples	INF001-E002b

4.6.1.4.3 Atribut: Nom

Identificador	ENT004-002
Nom formal	Nom
Nom curt	Nom
Domini	<p>Nom assignat a l'evidència dins el procés. La pauta per a la construcció del nom és la següent:</p> <ul style="list-style-type: none"> • Nom de la tipologia documental assignada en funció de la part declarativa del document • Objecte sobre el qual s'aplica la declaració del document
Tipus de dades	Alfanumèric
Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	<p>Aquest atribut equival a la propietat 3 Name del model AGRkMS.</p> <p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE3 – Nombre.</p> <p>Equival a l'atribut title que preveu la norma ISO 23081.</p> <p>Correspon a l'atribut genèric RiC-A28 Name.</p> <p>Correspon a l'element ISAD 3.1.2 Title.</p> <p>Correspon a l'element EAD <unittitle>.</p> <p>Correspon a l'atribut <DOC_NOMBRE> del model NEDA.</p> <p>DCAT proposa l'element de Dublin Core dct:title per indicar el nom del recurs que conté les dades.</p> <p>Aquest atribut és equivalent a la metadada Title del model ONDC.</p>
Exemples	Sol·licitud de beca per al menjador escolar

4.6.1.4.4 Atribut: Tipus

Identificador	ENT004-003
Nom formal	Tipus d'evidència
Nom curt	Tipus
Domini	Descripció del tipus genèric d'evidència. Determina si s'està descrivint un conjunt determinat i finit de dades agrupades en un mateix objecte, o bé un conjunt de dades dins un sistema d'informació o base de dades, de límits no determinables per cap altre mitjà que la mera enumeració de les dades.
Tipus de dades	Alfanumèric categòric
Mín. ocurrences	1
Màx. ocurrences	1
Regles de validació	L'atribut pot prendre dos valors: <ul style="list-style-type: none">• Document (entès en tant que document-objecte)• Dades (entès en tant que dades emmagatzemades directament en una base de dades)
Valor per defecte	No aplica.
Fonts	Correspon a l'atribut <DOC_TIPOSOPORTE> del model NEDA, però també es podria assimilar a l'atribut <DOC_TIPOCONT>. Equival, tot i que en una forma condensada únicament a dos possibles valors, a l'atribut RiC-A10 Content type . DCAT proposa l'element de Dublin Core dct:type per indicar el tipus de document o recurs. Per a fer-ho estableix l'obligatorietat d'utilitzar vocabularis controlats d'ús i governança consolidats. Aquest atribut també es pot considerar equivalent a la metadada Resource Type del model ONDC .
Exemples	Document

4.6.1.4.5 Atribut: Caràcter

Identificador	ENT004-004
Nom formal	Caràcter de l'evidència
Nom curt	Caràcter
Domini	Descripció de l'autenticitat de l'evidència. En aquest sentit, hi ha tres caràcters possibles: <ul style="list-style-type: none">• Original: el document emana directament del seu autor (productor), que en deixa constància de l'autoria mitjançant algun mitjà d'autenticació.• Còpia autèntica: el document emana d'un agent intermedi o de l'agent receptor, el qual ha marcat una còpia d'un document original com autèntica un cop haver-la acarat amb l'original i haver comprovat la identitat d'ambdós continguts.

	<ul style="list-style-type: none"> • Còpia simple: es desconeix de qui emana el document, en no constar-hi cap marca o signatura que n'acrediti l'autenticitat. El document reproduceix en aparença la forma externa d'un altre document, però no proporciona mitjans per verificar-ne la coincidència real.
Tipus de dades	Alfanumèric categòric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	<p>L'atribut pot prendre tres valors:</p> <ul style="list-style-type: none"> • Original • Còpia autèntica • Còpia simple
Valor per defecte	No aplica.
Fonts	<p>Correspon a l'atribut RiC-A39 State, el qual determina mitjançant una llista de valors l'estat d'elaboració d'una unitat documental. En aquest model es recullen només estats finals, tot i que RiC contempla també estats previs (per exemple "esborrany"). RiC també ho contempla en forma de relació, establint el vincle entre documents amb les relacions RiC-R010 Record is original of Record i inversa, RiC-R011 Record is draft of Record i inversa, i RiC-R012 Record is copy of Record i inversa.</p> <p>ISAD disposa de dos elements que fan referència a aquest atribut, malgrat que no hi coincideixen en la significació. Es tracta de l'element 3.5.1 Existence and location of originals, per a les còpies, i de l'element 3.5.2 Existence and location of copies, per als originals.</p> <p>DCAT, en fer referència a les relacions entre recursos (<i>datasets</i>) que inclouen un rol (per exemple "és esborrany de", "és còpia de", etc.) proposa l'element dcat:qualifiedRelation, combinat amb l'ús de vocabularis controlats.</p>
Exemples	Original

4.6.1.4.6 Atribut: Format

Identificador	ENT004-005
Nom formal	Format
Nom curt	Format
Domini	Format físic-lògic del document. Fa referència tant als suports físics, com al format en el cas de documents electrònics.
Tipus de dades	Alfanumèric categòric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	L'atribut es valida contra una llista de valors creada a partir dels formats detectats. En aquesta llista es contemplen les següents casuístiques:

	<ul style="list-style-type: none"> • Documents en suports no digitals: s'informa el suport. • Documents-objecte digitals: s'informa el format de fitxer amb el nom i identificador del registre de formats PRONOM (<i>PRONOM Unique Identifier</i>, o PUID). • Dades en aplicacions informàtiques: s'informa el nom de la tecnologia o plataforma on s'emmagatzemen les dades.
Valor per defecte	No aplica.
Fonts	<p>ISAD no contempla un element específic per a descriure el format, però permet incloure informació relativa al format en dos elements: 3.1.5 Extent and medium of the unit of description i 3.4.4 Physical characteristics and technical requirements.</p> <p>EAD disposa de l'element <phystech> equivalent a aquesta entitat.</p> <p>Aquest atribut equival a la propietat 19 Format del model AGRkMS.</p> <p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE14.1 - Formato.</p> <p>Equival a l'atribut technical environment que preveu la norma ISO 23081.</p> <p>Equival indirectament a l'atribut <DOC_TIPOMED> del model NEDA.</p> <p>RiC disposa de diferents atributs per a descriure el format d'una instància d'unitat documental. Permet definir el suport amb RiC-A50 Carrier type, la tècnica de creació amb RiC-A33 Production technique, la forma de codificar-ne la representació amb RiC-A37 Representation type, o fins i tot amb l'atribut RiC-A17 Documentary Form Type, en el moment que esmenta com a forma documental el correu electrònic. Malgrat això, per indicar formats de fitxer concrets probablement caldrà recórrer a l'atribut RiC-A41 Technical characteristics, el qual, malgrat sigui aplicable només a l'entitat RiC-E13 Mechanism, inclou el possible esment d'elements de programari.</p> <p>DCAT proposa l'element de Dublin Core dct:format per descriure el format i característiques físiques del recurs o del document. Per estandarditzar la definició dels formats DCAT proposa un sub-element de dct:format, l'element dc:mediaType, que preceptivament cal informar amb els tipus definits per la IANA.²¹⁵</p> <p>Aquest atribut és equivalent a la metadada Format del model ONDC.</p>
Exemples	Paper

²¹⁵ El registre de formats de la IANA (*Internet Assigned Numbers Authority*), disponible a <https://www.iana.org/assignments/media-types/media-types.xhtml>.

4.6.1.4.7 Atribut: Autenticació

Identificador	ENT004-006
Nom formal	Sistema d'autenticació
Nom curt	Autenticació
Domini	Sistema utilitzat per autenticar l'evidència. En cas d'existir diferents sistemes d'autenticació, es farà constar només el de més alt nivell de seguretat.
Tipus de dades	Alfanumèric categòric
Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	<p>L'atribut pot prendre els següents valors:</p> <ul style="list-style-type: none"> • Sense autenticació • Signatura autògrafa (realitzada de forma manual) • Segellat manual (realitzat de forma manual amb suport mecànic) • Signatura electrònica (signatura basada en certificat electrònic, que pot ser avançada o reconeguda/qualificada) • Segellat electrònic (signatura basada en certificat electrònic, que pot ser avançada o reconeguda/qualificada) • Cadena de blocs (signatura vinculada a l'ús de <i>Blockchain</i>) • Codi de verificació (signatura vinculada a la possibilitat d'accedir a l'original a través d'un codi segur de verificació) • Pista d'auditoria (autenticació a través del registre d'accions realitzades sobre l'evidència per part del sistema d'informació que la conté) • Metadades (autenticació a través de les metadades relacionades amb l'origen i valor de l'evidència)
Valor per defecte	No aplica.
Fonts	<p>Aquest atribut equival a la propietat 22 Integrity check del model AGRkMS.</p> <p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE17 - Firma. I també, però ja no de forma directa, amb la metadada eEMGDE16 - Verificación de integridad.</p> <p>Equival a l'atribut integrity que preveu la norma ISO 23081.</p> <p>Equivalent a l'atribut RiC-A03 Authenticity note, que descriu de quina forma s'autentifica la unitat de descripció. Però no a l'atribut RiC-A24 Integrity, que fa referència només a instàncies concretes, element que no es recull en aquest model.</p>
Exemples	Signatura autògrafa

4.6.1.4.8 Atribut: Registre d'entrada

Identificador	ENT004-007
Nom formal	Registre d'entrada
Nom curt	Registre d'entrada
Domini	<p>S'indica en aquest atribut l'existència d'un registre d'entrada en l'evidència. S'entén per registre d'entrada el que estableix la Llei 39/2015 del procediment administratiu com a sistema habitual de constància de recepció a l'Administració pública,²¹⁶ però també es considera registre d'entrada qualsevol altre sistema que permeti acreditar de forma fefaent la data i hora de recepció d'un document o agregat de dades en una organització.</p> <p>Aquest atribut s'ha inclòs en el model com a resultat de necessitats específiques de gestió de la informació del Consorci d'Educació de Barcelona. No es considera determinant als efectes d'anàlisi de les dades recollides.</p>
Tipus de dades	Booleà
Mín. ocurrences	1
Màx. ocurrences	1
Regles de validació	Pot prendre dos valors: <ul style="list-style-type: none">• Sí• No
Valor per defecte	"No"
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	Sí

4.6.1.4.9 Atribut: Registre de sortida

Identificador	ENT004-008
Nom formal	Registre de sortida
Nom curt	Registre de sortida
Domini	<p>S'indica en aquest atribut l'existència d'un registre de sortida en l'evidència. S'entén per registre de sortida el que estableix la Llei 39/2015 del procediment administratiu com a sistema habitual de constància de tramesa d'un document des de l'Administració pública cap a qualsevol altre destinatari, però també es considera registre de sortida qualsevol altre sistema que permeti acreditar de forma fefaent la data i hora de tramesa d'un document o agregat de dades en una organització.</p> <p>Aquest atribut s'ha inclòs en el model com a resultat de necessitats específiques de gestió de la informació del Consorci d'Educació de</p>

²¹⁶ Ley 39/2015, de 1 de octubre, del Procedimiento Administrativo Común de las Administraciones Públicas.

	Barcelona. No es considera determinant als efectes d'anàlisi de les dades recollides.
Tipus de dades	Booleà
Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	Pot prendre dos valors: <ul style="list-style-type: none"> • Sí • No
Valor per defecte	"No"
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	Sí

4.6.1.4.10 Atribut: *Essencial*

Identificador	ENT004-009
Nom formal	Essencial
Nom curt	Essencial
Domini	Aquest atribut recull una característica contextual de tota evidència, que és el seu caràcter d'essencial. S'entén per evidència essencial dins un conjunt documental o d'informació aquella evidència que conté informació substantiva o determinant per a la significació del conjunt. Això vol dir: <ul style="list-style-type: none"> • Que l'evidència conté informació (dades) que no es troben en cap altra evidència del mateix conjunt. • Que aquesta informació documenta parts de l'acció que ha donat origen al conjunt sense les quals no es pot entendre el global de l'acció, o bé aquesta queda incompleta. <p>Si ho identifiquem amb l'anàlisi de processos, diríem que una evidència essencial és aquella que documenta una acció ubicada en el camí crític del procés.</p>
Tipus de dades	Booleà
Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	Pot prendre dos valors: <ul style="list-style-type: none"> • Sí • No
Valor per defecte	"No"
Fonts	El model e-EMGDE contempla per aquest atribut la metadada eEMGDE13.4 - Documento esencial .
Exemples	Sí

4.6.1.4.11 Atribut: Data creació

Identificador	ENT004-010
Nom formal	Data de creació
Nom curt	Data creació
Domini	La data de creació d'una evidència és la data en què aquesta evidència ha estat identificada i descrita dins el procés.
Tipus de dades	Data
Mín. ocurrences	1
Màx. ocurrences	1
Regles de validació	Data vàlida.
Valor per defecte	No aplica.
Fonts	<p>Aquest atribut equival a la propietat 4 Date range del model AGRkMS.</p> <p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE4.1 - Fecha inicio.</p> <p>Correspon a l'element ISAD 3.1.3 Date(s).</p> <p>Correspon a l'element EAD <unitdate>.</p> <p>Correspon a l'atribut <DOC_FECHA> del model NEDA.</p> <p>Equival a l'atribut RiC-A29 Normalized date dins l'entitat RiC-E20 Date range. Vegeu nota explicativa a l'apartat Fonts de l'atribut ENT001-006.</p>
Exemples	12 de gener de 2008.

4.6.1.4.12 Atribut: Data supressió

Identificador	ENT004-011
Nom formal	Data de supressió
Nom curt	Data supressió
Domini	La data de supressió és la data en què una evidència deixa d'aparèixer en el curs d'un procés.
Tipus de dades	Data
Mín. ocurrences	1
Màx. ocurrences	1
Regles de validació	Data vàlida.
Valor per defecte	No aplica.
Fonts	<p>Aquest atribut equival a la propietat 4 Date range del model AGRkMS.</p> <p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE4.2 - Fecha fin.</p> <p>Correspon a l'element ISAD 3.1.3 Date(s).</p>

	<p>Correspon a l'element EAD <unitdate>.</p> <p>Correspon a l'atribut <DOC_FECHA> del model NEDA.</p> <p>Equival a l'atribut RiC-A29 Normalized date dins l'entitat RiC-E20 Date range. Vegeu nota explicativa a l'apartat Fonts de l'atribut ENT001-006.</p>
Exemples	31 de desembre de 2021.

4.6.1.5 Entitat: Model de document

4.6.1.5.1 Descripció de l'entitat

Identificador	ENT005
Nom formal	Model de document
Nom curt	Model de document
Domini	<p>Aquesta entitat descriu els diferents models o plantilles que defineixen l'estructura d'evidències que han estat detectades dins el procés. Habitualment s'implementa en forma de formularis, models de documents, o pantalles d'entrada de dades. Mantenim la denominació "Model de document" per facilitar la seva comprensió, atès que és la denominació més usual, tot i que en el cas de pantalles d'entrada de dades s'utilitzarà habitualment el terme "Formulari".</p> <p>Habitualment els models de document, especialment quan van vinculats a convocatòries temporals, són modificats de forma no substancial per a cada convocatòria. Aquests canvis lleus donen origen a noves versions, be que mantenint l'essència del model de document original. Tot i així, cada versió anirà associada a un objecte digital diferent. Per aquest motiu, aquesta entitat recollirà com una nova instància cadascuna de les versions dels models de document descrits.</p>
Atributs	ENT005-001 Identificador ENT005-002 Versió ENT005-003 Nom ENT005-004 Tipus ENT005-005 Format ENT005-006 Data creació ENT005-007 Data supressió
Relacions	REL021 Formalitza la forma externa d'una o més evidències
Fonts	Cap model dels analitzats esmenta explícitament aquest concepte com entitat. De totes maneres, en ser aquesta entitat derivada de l'entitat ENT004 Evidència , li assignem en alguns casos atributs de determinades fonts que podrien ser-li aplicables en cas d'estar reconeguda com entitat en els diferents models.

4.6.1.5.2 Atribut: Identificador

Identificador	ENT005-001
Nom formal	Identificador
Nom curt	ID
Domini	Codi que identifica de forma unívoca cada versió de cada model de document.
Tipus de dades	Alfanumèric
Mín. ocurrencies	1

Màx. ocurrencies	1
Regles de validació	L'identificador de cada model de document s'assigna com un codi qualificat, que es compon dels següents elements: <ul style="list-style-type: none"> • Una lletra en majúscula, que indica el tipus de model • L'identificador del model de document, un número seqüencial de quatre dígit • Guió • Lletra V (indicativa de versió) • Número seqüencial de dos dígit, reiniciat per a cada model de document • Guió • Dos darrers dígit de l'any de publicació de la versió
Valor per defecte	No aplica.
Fonts	Equival a l'atribut identfier que preveu la norma ISO 23081 . Correspon a l'atribut genèric RiC-A22 Identifier .
Exemples	A0002-V01-18

4.6.1.5.3 Atribut: Versió

Identificador	ENT005-002
Nom formal	Identificador de versió
Nom curt	Versió
Domini	Codi de identifica de forma unívoca cadascuna de les versions d'un model de document. Aquesta informació es recull de forma reiterativa dins el codi del model de document.
Tipus de dades	Alfanumèric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	2

4.6.1.5.4 Atribut: Nom

Identificador	ENT005-003
Nom formal	Nom
Nom curt	Nom
Domini	Denominació formal del model de document. Habitualment la denominació es manté entre diferents versions del document.
Tipus de dades	Alfanumèric
Mín. ocurrencies	1
Màx. ocurrencies	1

Regles de validació	No
Valor per defecte	No aplica.
Fonts	Equival a l'atribut title que preveu la norma ISO 23081 . Correspon a l'atribut genèric RiC-A28 Name .
Exemples	Sol·licitud d'actuació de l'equip d'atenció i orientació psicopedagògica

4.6.1.5.5 Atribut: Tipus

Identificador	ENT005-004
Nom formal	Tipus de document
Nom curt	Tipus
Domini	El tipus descriu la categoria a la qual pertany el model de document en funció de la seva part declarativa. Correspon al que habitualment es coneix com a tipologia documental, i és habitualment recollida en un quadre de tipologies documentals com a part dels instruments descriptius d'un sistema de gestió documental. El sistema de classificació utilitzat en aquest atribut es desplega en diferents nivells de concreció, de manera que permet assignar sempre una classificació, més genèrica o més específica en funció del context d'aplicació del model de document.
Tipus de dades	Alfanumèric categòric
Mín. ocurrences	1
Màx. ocurrences	1
Regles de validació	Per classificar aquest atribut s'ha utilitzat el Quadre de Tipus Documentals corporatiu de la Generalitat de Catalunya, ²¹⁷ incloent el nom i, entre parèntesi, el codi assignat.
Valor per defecte	No aplica.
Fonts	Aquest atribut equival a la propietat 24 Document form del model AGRkMS . Equival a l'atribut documentary form que preveu la norma ISO 23081 . El model e-EMGDE contempla per aquest atribut la metadada eEMGDE18 - Tipo documental . Correspon a l'atribut <DOC_TIPODOC> del model NEDA . La correspondència d'aquest atribut amb l'atribut RiC-A17 Documentary Form Type de RiC és exacta, doncs descriu el concepte clàssic de tipus o tipologia documental, present en gairebé totes les tradicions arxivístiques.

²¹⁷

Vegeu <https://cultura.gencat.cat/ca/temes/arxius-i-gestio-documental/gestio-documental/gdgencat/aresta/protocol/instruments/qdtdoc/>.

Exemples	Sol·licitud (TD14-010)

4.6.1.5.6 Atribut: *Format*

Identificador	ENT005-005
Nom formal	Format
Nom curt	Format
Domini	Format físic-lògic del document. Fa referència tant als suports físics, com al format en el cas de documents electrònics.
Tipus de dades	Alfanumèric categòric
Mín. ocurrences	1
Màx. ocurrences	1
Regles de validació	<p>L'atribut es valida contra una llista de valors creada a partir dels formats detectats. En aquesta llista es contemplen les següents casuístiques:</p> <ul style="list-style-type: none"> • Formularis o models de document en suports no digitals: s'informa el suport. • Formularis o models de document estàtics o en forma de document-objecte: s'informa el format de fitxer amb el nom i identificador del registre de formats PRONOM (<i>PRONOM Unique Identifier</i>, o PUID). • Formularis interactius o vinculats a aplicacions informàtiques: s'informa el nom de la tecnologia o plataforma que s'ha utilitzat per a desenvolupar el formulari.
Valor per defecte	No aplica.
Fonts	<p>ISAD no contempla un element específic per a descriure el format, però permet incloure informació relativa al format en dos elements: 3.1.5 Extent and medium of the unit of description i 3.4.4 Physical characteristics and technical requirements.</p> <p>Aquest atribut equival a la propietat 19 Format del model AGRkMS.</p> <p>Equival a l'atribut technical environment que preveu la norma ISO 23081.</p> <p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE14.1 - Formato.</p> <p>RiC disposa de diferents atributs per a descriure el format d'una instància d'unitat documental. Permet definir el suport amb RiC-A50 Carrier type, la tècnica de creació amb RiC-A33 Production technique, la forma de codificar-ne la representació amb RiC-A37 Representation type, o fins i tot amb l'atribut RiC-A17 Documentary Form Type, en el moment que esmenta com a forma documental el correu electrònic. Malgrat això, per indicar formats de fitxer concrets probablement caldrà recórrer a l'atribut RiC-A41 Technical characteristics, el qual, malgrat sigui aplicable només a l'entitat RiC-</p>

	<p>E13 Mechanism, inclou el possible esment d'elements de programari.</p> <p>DCAT proposa l'element de Dublin Core dct:format per descriure el format i característiques físiques del recurs o del document. Per estandarditzar la definició dels formats DCAT proposa un subelement de dct:format, l'element dcat:mediaType, que preceptivament cal informar amb els tipus definits per la IANA.²¹⁸</p> <p>Aquest atribut és equivalent a la metadada Format del model ONDC.</p>
Exemples	Paper

4.6.1.5.7 Atribut: Data creació

Identificador	ENT005-006
Nom formal	Data de creació
Nom curt	Data creació
Domini	La data de creació d'un model de document és la data en què s'inicia el desenvolupament del formulari o model de document.
Tipus de dades	Data
Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	Data vàlida.
Valor per defecte	No aplica.
Fonts	Equival a l'atribut RiC-A29 Normalized date dins l'entitat RiC-E20 Date range . Vegeu nota explicativa a l'apartat Fonts de l'atribut ENT001-006 .
Exemples	12 de gener de 2008.

4.6.1.5.8 Atribut: Data supressió

Identificador	ENT005-007
Nom formal	Data de supressió
Nom curt	Data supressió
Domini	La data de supressió és la data en què un model de document deixa d'estar actiu i utilitzar-se dins l'organització. En els models digitals és la data en què es despublica.
Tipus de dades	Data
Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	Data vàlida.
Valor per defecte	No aplica.

²¹⁸ El registre de formats de la IANA (*Internet Assigned Numbers Authority*), disponible a <https://www.iana.org/assignments/media-types/media-types.xhtml>.

Fonts	Equival a l'atribut RiC-A29 Normalized date dins l'entitat RiC-E20 Date range . Vegeu nota explicativa a l'apartat Fonts de l'atribut ENT001-006 .
Exemples	31 de desembre de 2021.

4.6.1.6 Entitat: Dada elemental

4.6.1.6.1 Descripció de l'entitat

Identificador	ENT006
Nom formal	Dada elemental
Nom curt	Dada
Domini	Element d'informació mínim amb unitat de significació. La dada elemental pot estar composta de dades de menor dimensió (dades atòmiques). Però es considera dada elemental quan té una unitat de significació, i pot ser reconeixible a partir d'aquesta significació. La dada elemental pot ser element d'anàlisi en un sistema d'informació.
Atributs	ENT006-001 Identificador ENT006-002 Nom ENT006-003 Tipus de dada ENT006-004 Descripció ENT006-005 Regles de validació ENT006-006 Nivell seguretat RGPD ENT006-007 Disponible per interoperabilitat externa
Relacions	REL020 Apareix a una o més evidències REL022 Pot ser traslladada/migrada per una o més interfícies REL023 Pot estar continguda a un o més sistemes d'informació REL026 Descriu un objecte REL027 Pot aparèixer a un o més quadres de comandament
Fonts	<p>Malgrat que no ho desenvolupa ni ho descriu explícitament com a tal, el model NEDA inclou una entitat que contempla el concepte de dada elemental, concretament en el subtipus Componente documental de l'entitat Documentos de archivo.</p> <p>Es podria trobar una correspondència entre la dada, tal com s'identifica en aquest model, i l'entitat RiC-E05 Record part del model RiC. D'acord amb la definició que RiC fa del concepte de <i>Record part</i> ("part d'un document amb un contingut d'informació discreta que contribueix a la completesa física o intel·lectual del document") podem posar-lo en relació amb el concepte de dada elemental que s'utilitza en aquest model.</p> <p>Una dels orígens on és més evident identificar dades és en els models de dades d'una base de dades relacional. Però malgrat aquesta facilitat, la relació no és directa. Per aquest motiu, s'ha considerat les possibles equivalències entre l'entitat dada elemental i els nivells de metadescripció de SIARD corresponents a column (requeriment M_5.6-1) i field (requeriment M_5.7-1).</p> <p>El model DCAT no contempla específicament el concepte de dada sense estar vinculada a cap conjunt de dades, però implícitament la contempla en proporcionar la possibilitat de descriure l'estructura d'un dcat:Dataset. Ho vinculem per tant per aquest punt.</p>

	El mateix s'esdevé amb ONDC . Malgrat el concepte de <i>data asset</i> és equiparable al de conjunt de dades (l'element Resource type no el contempla com a tal), no volem descartar la seva aplicació a l'entitat dada elemental d'aques model.
--	--

4.6.1.6.2 Atribut: Identificador

Identificador	ENT006-001
Nom formal	Identificador
Nom curt	Id
Domini	Identificador únic de la dada.
Tipus de dades	Alfanumèric
Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	Correspon a l'atribut genèric RiC-A22 Identifier . És equivalent a l'element dct:identificer del model DCAT . Pot considerar-se equivalent a l'element Identifier del model ONDC .
Exemples	00005

4.6.1.6.3 Atribut: Nom

Identificador	ENT006-002
Nom formal	Nom
Nom curt	Nom
Domini	Nom assignat a la dada. Es tracta d'un nom curt, que descriu el contingut de la dada.
Tipus de dades	Alfanumèric
Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	Correspon a l'atribut genèric RiC-A28 Name . Es podrien trobar equivalències amb la metadada name dels nivells de descripció column (M_5.6-1) i field (M_5.7-1) de SIARD . És equivalent a l'element dct:title del model DCAT . Pot considerar-se equivalent a l'element Title del model ONDC .
Exemples	Adreça

4.6.1.6.4 Atribut: Tipus de dada

Identificador	ENT006-003
Nom formal	Tipus de dada
Nom curt	Tipus
Domini	Tipus d'informació que conté la dada.
Tipus de dades	Alfanumèric categòric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	Els valors s'assignen a partir d'una taula incremental de tipus de dades, que conté els tipus més habituals: dades identificatives, dades geogràfiques, etc.
Valor per defecte	No aplica.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	Acadèmiques

4.6.1.6.5 Atribut: Descripció

Identificador	ENT006-004
Nom formal	Descripció
Nom curt	Descripció
Domini	Descripció textual de la significació de la dada i de la seva relació amb l'objecte descrit o caracteritzat.
Tipus de dades	Alfanumèric descriptiu
Mín. ocurrencies	0
Màx. ocurrencies	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	Podem trobar equivalències amb la metadada description dels nivells de descripció column (M_5.6-1) i field (M_5.7-1) de SIARD . És equivalent a l'element dct:description del model DCAT . Pot considerar-se equivalent a l'element Description del model ONDC .
Exemples	L'Eina de seguiment de les actuacions d'orientació del Pla Jove sobre l'alumnat conté la situació actual del jove que disposa del Servei d'Orientació. Conté els estudis que es cursen, l'especialitat, el centre, el nivell (0, 1, 2 (CFGM, BATX) 3 (PACFGS, CFGS).

4.6.1.6.6 Atribut: Regles de validació

Identificador	ENT006-005
Nom formal	Regles de validació
Nom curt	Regles de validació

Domini	Regles que determinen la forma com s'informen els valors d'aquella dada, no en un sistema en concret sinó de forma general.
Tipus de dades	Alfanumèric
Mín. ocurrencies	0
Màx. ocurrencies	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	<p>Taula de motius de necessitats específiques de suport educatiu:</p> <ul style="list-style-type: none"> • NESE derivades d'altres capacitats • NESE derivades de discapacitat auditiva • NESE derivades de discapacitat visual • NESE derivades de discapacitat intel·lectual • NESE derivades de discapacitat física • NESE derivades de pluridiscapacitat • NESE derivades de trastorns de l'espectre de l'autisme • NESE derivades de trastorn greu de conducta • NESE derivades de trastorn mental • NESE derivades de trastorns de l'aprenentatge i la comunicació • NESE derivades de situació de malaltia • NESE derivades de desavantatge educatiu

4.6.1.6.7 Atribut: Nivell de seguretat RGPD

Identificador	ENT006-006
Nom formal	Nivell de seguretat de protecció de dades personals
Nom curt	Nivell de seguretat RGPD ²¹⁹
Domini	<p>Nivell de seguretat de la dada en relació amb la protecció de dades de caràcter personal. Només s'aplica a les dades que tenen per objecte una persona, i s'indica si aquestes dades tenen la consideració de dada sensible o que ha de ser objecte d'especial protecció.</p> <p>Aquest atribut s'ha inclòs en el model com a resultat de necessitats específiques de gestió de la informació del Consorci d'Educació de Barcelona. No es considera determinant als efectes d'anàlisi de les dades recollides.</p>
Tipus de dades	Alfanumèric categòric
Mín. ocurrencies	1
Màx. ocurrencies	1

²¹⁹ Inicials del Reglament (UE) 2016/679 del Parlament i del Consell, de 27 d'abril de 2016, relatiu a la protecció de les persones físiques pel que fa al tractament de dades personals i a la lliure circulació d'aquestes dades i pel qual es deroga la Directiva 95/46/CE (Reglament general de protecció de dades) (DOUE 4.5.2016). Aquest reglament és la norma bàsica que regula a escala europea la protecció de dades de caràcter personal.

Regles de validació	Aquest atribut pot prendre dos valors: <ul style="list-style-type: none"> • Sensible • No sensible
Valor per defecte	No aplica.
Fonts	El model e-EMGDE contempla per aquest atribut la metadada eEMGDE8.4 - Sensibilidad datos de carácter personal . I també amb la metadada eEMGDE8.6 - Nivel de confidencialidad de la información . Es pot posar en relació amb l'element dct:accessRights del model DCAT . Aquest atribut és equivalent a la metadada Security Classification del model ONDC , i també a la metadada Sensitive Data .
Exemples	Sensible

4.6.1.6.8 Atribut: *Disponible per interoperabilitat externa*

Identificador	ENT006-007
Nom formal	Disponible per interoperabilitat externa
Nom curt	Interoperable
Domini	Indica si la dada pot ser objecte d'intercanvi amb altres organitzacions a través d'algun sistema d'informació. S'indica en aquest camp el codi del sistema o sistemes d'informació a través dels quals pot ser interoperable. Aquest atribut s'ha inclòs en el model com a resultat de necessitats específiques de gestió de la informació del Consorci d'Educació de Barcelona. No es considera determinant als efectes d'anàlisi de les dades recollides.
Tipus de dades	Alfanumèric categòric
Mín. ocurrencies	0
Màx. ocurrencies	1
Regles de validació	Aquest atribut agafa els seus valors de la relació de sistemes d'informació, a través del codi de sistema.
Valor per defecte	No aplica.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	SI02.014

4.6.1.7 Entitat: Sèrie documental

4.6.1.7.1 Descripció de l'entitat

Identificador	ENT007
Nom formal	Sèrie documental
Nom curt	Sèrie documental
Domini	<p>S'entén per sèrie documental el conjunt de tots els documents o evidències creats o rebuts per un agent (productor) com a conseqüència de la seva participació en un procés de treball (la majoria de definicions parlen de “resultat de funcions i activitats”), i que esdevenen una prova de processos de treball.</p> <p>Les sèries documentals són el resultat de la identificació realitzada per un professional de la gestió documental sobre el substrat evidencial d'un procés de treball.</p> <p>Per una definició detallada del concepte de sèrie documental, vegeu l'apartat 2.3.1 i 2.3.2 de l'estat de la qüestió (capítol 2).</p>
Atributs	<p>ENT007-001 Identificador</p> <p>ENT007-002 Nom</p> <p>ENT007-003 Criteri de nominalització</p> <p>ENT007-004 Caràcter</p> <p>ENT007-005 Descripció</p> <p>ENT007-006 Vital</p> <p>ENT007-007 100% Digital</p> <p>ENT007-008 Data creació</p> <p>ENT007-009 Data supressió</p>
Relacions	<p>REL004 És produïda per un agent</p> <p>REL012 Pot estar relacionada amb una o més sèries documentals</p> <p>REL013 És el resultat d'un procés</p> <p>REL014 És regulada per una o més regles d'accés</p> <p>REL015 És regulada per una o més regles de conservació</p> <p>REL018 Conté una o més evidències</p>
Fonts	<p>ISAD, la norma de descripció arxivística de referència, contempla de forma explícita l'entitat sèrie documental, juntament amb altres nivells d'agregació arxivístics (fons, subfons, subsèrie i expedient). Pel seu caràcter multinivell, l'entitat sèrie documental es fa explícita a través de l'element 3.1.4 Level of description.</p> <p>EAD també contempla la sèrie documental de forma explícita, citada com un dels nivells de descripció (atribut level) de l'element <archdesc>, que pot prendre els valors: <i>class, collection, file, fons, item, otherlevel, recordgrp, series, subfons, subgrp, subseries</i>.</p> <p>MoReq, en un intent de ser neutral en relació amb les diferents tradicions arxivístiques, no recull explícitament el concepte de sèrie ni els altres nivells de descripció ISAD, sinó que utilitza el concepte</p>

	<p>genèric d'agrupació documental, recollit a l'entitat E14.2.1 Aggregation.</p> <p>La norma ISO 23081 contempla Records com una de les cinc entitats principals del seu model, en un concepte ampli que també n'inclou les agrupacions.</p> <p>El model AGRkMS inclou l'entitat record, que defineix com: “[...] <i>A record may comprise an electronic or paper-based document or group of aggregated documents</i>”. A partir de l'esment del concepte d'agregació es podria inferir que pot incloure agrupacions documentals superiors a l'expedient, i en concret la sèrie documental, però no en fa cap esment específic. En el nostre cas no ho hem considerat equivalent a sèrie documental.</p> <p>Sí que ho és, en canvi, en el model e-EMGDE, atès que quan desenvolupa les diferents possibles agrupacions de l'entitat Documento (que anomena categories), enumera els nivells arxivístics de grup de fons, fons, sèrie (documental), agregació i expedient. Per tant la sèrie documental queda també inclosa en l'entitat Documento de e-EMGDE.</p> <p>El model NEDA contempla l'entitat Documentos de archivo, la qual subdivideix en deu subtipus: Grupo de fondos, Fondo, División de fondo o de grupo de fondo, Serie, Subserie, Fracción de serie/subserie, Unidad documental, Colección, División de colección i Componente documental. En aquest cas l'entitat correspondria estrictament al subtipus Serie, tot i que, en la significació que li donem en aquest model, podria englobar tots els nivells d'agrupació documental arxivístics, que són Serie, Subserie, Fracción de serie/subserie, Colección i División de colección.</p> <p>RiC no recull el concepte de sèrie documental a nivell d'entitat. Ho argumenta pel fet que en la disciplina arxivística no existeix un acord unànimе sobre el concepte de sèrie, atès que sovint s'ha treballat amb diferents nivells d'agrupació, no sempre homogenis ni en la concreció ni en el criteri d'agrupació (fons, subfons, col·leccions, sèries, subsèries, etc.), la qual cosa ha generat ambigüitat en moltes descripcions arxivístiques. Per aquest motiu, RiC proposa una única entitat neutra, que anomena RiC-E03 Record set, la qual depèn de RiC-E02 Record resource, i que es pot subdividir en diferents nivells gràcies al caràcter multidimensional de l'esquema RiC. El nivell sèrie el recull en l'atribut RiC-A36 Record set type, com a categorització de l'entitat RiC-E03 Record set.</p> <p>La sèrie documental, com a resultat d'un procés de dadificació, es materialitza en una matriu de dades o variables d'informació. D'acord amb això, podríem assimilar la sèrie documental al concepte de vista d'una base de dades relacional, que trobem recollit en l'element view (requeriment M_5.14-1) del model de metadades</p>
--	---

	<p>SIARD. En l'apartat de conclusions desenvolupem amb més detall aquesta possible assimilació.</p> <p>DCAT, quan planteja la descripció de conjunts de dades (<i>datasets</i>), distingeix entre una versió abstracta del <i>dataset</i> i les seves diferents representacions tecnològiques, que anomena <i>distributions</i>. Aquesta versió abstracta del <i>dataset</i> correspon a la classe dc:Resource, que és a la seva vegada la superclasse de dc:Dataset i dc:DataService. De la mateixa manera que ho hem indicat amb SIARD, aquesta versió abstracta del <i>dataset</i>, que correspon a un conjunt o matriu de dades, seria assimilable també a una sèrie documental dadificada. Tot depèn, en tot cas, de la seva significació (funcional, documental o tecnològica).</p> <p>L'esquema ONDC permet descriure el que anomena actius d'informació ("<i>information assets</i>") als efectes d'elaboració d'un catàleg de dades. Dins aquesta categoria hi engloba documents-objecte, imatges, conjunts de dades estructurades ("<i>datasets</i>"), aplicacions informàtiques, etc. Però la correspondència que en proposen els Arxius Nacionals d' Austràlia ho equipara especialment a l'entitat "<i>Record</i>", que inclou tant els documents com les seves agrupacions.</p>
--	--

4.6.1.7.2 Atribut: Identificador

Identificador	ENT007-001
Nom formal	Identificador
Nom curt	Id
Domini	Codi que identifica de forma unívoca una sèrie documental dins el conjunt de totes les sèries d'una organització.
Tipus de dades	Alfanumèric
Mín. ocurrences	1
Màx. ocurrences	1
Regles de validació	L'identificador de la sèrie documental és un codi qualificat, que es compon del codi del procés i d'un número continu de tres dígits, separat per un guió.
Valor per defecte	No aplica.
Fonts	<p>Correspon a l'element ISAD 3.1.1 Reference code. Coincideix de forma exacta amb l'ús que se'n fa en aquest model, atès que no és redundant amb el codi de classificació funcional (en el nostre cas el codi de procés).</p> <p>EAD disposa de diferents elements per a descriure l'identificador d'una agrupació documental: <recordid> i <unitid>.</p> <p>El model MoReq recull aquest atribut en la metadada M14.4.100 System Identifier.</p> <p>Equival a l'atribut identif que preveu la norma ISO 23081.</p>

	<p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE22 – Clasificación, però només quan s'utilitza en modalitat monoentitat. També es pot considerar equivalent la metadada eEMGDE2 – Identificador.</p> <p>Correspon a l'atribut <DOC_ID> del model NEDA.</p> <p>Correspon a l'atribut genèric RiC-A22 Identifier. En aquest cas, a més, aquest atribut també coincideix amb l'atribut RiC-A07 Classification. Malgrat RiC especifica que l'identificador i la classificació s'han de considerar dos elements diferents, no exclou la possibilitat que formalment coincideixen si l'identificador s'utilitza al mateix temps com a codi de classificació d'unitats documentals.</p> <p>Aquest atribut és equivalent a la metadada Identifier del model ONDC.</p>
Exemples	ORG003-001

4.6.1.7.3 Atribut: Nom

Identificador	ENT007-002
Nom formal	Nom
Nom curt	Nom
Domini	<p>Correspon al títol assignat a la sèrie documental pel professional de la gestió documental, d'acord amb la significació de la sèrie documental dins un context funcional determinat. El títol es compon habitualment de la concatenació del nom del procés i del nom del productor o agent responsable d'aquella sèrie, atès que una sèrie és el resultat de la participació d'un determinat agent en un determinat procés. A vegades, però, per motius pràctics o de significació comunament reconeguda es poden utilitzar altres títols, sovint descriptius del tipus de documentació que compona la sèrie, o de la matèria sobre la qual aquella sèrie conté informació.</p> <p>En alguns casos, especialment quan la sèrie ja té un nom consagrat per l'ús, es pot diferenciar entre un títol propi (el títol assignat pel professional de la gestió documental) i un títol atribuït (el títol utilitzat pels usuaris habituals de la sèrie).</p>
Tipus de dades	Alfanumèric
Mín. ocurrences	1
Màx. ocurrences	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	<p>Correspon a l'element ISAD 3.1.2 Title.</p> <p>Correspon a l'element EAD <unittitle>.</p>

	<p>El model MoReq recull aquest atribut en la metadada M14.4.104 Title.</p> <p>Equival a l'atribut title que preveu la norma ISO 23081.</p> <p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE3 – Nombre.</p> <p>Correspon a l'atribut <DOC_NOMBRE> del model NEDA.</p> <p>Correspon a l'atribut genèric RiC-A28 Name.</p> <p>D'acord amb l'assimilació feta a nivell d'entitat, podríem equiparar aquest atribut a la metadada M_5.14-1 name del model SIARD.</p> <p>Aquest atribut és equivalent a la metadada Title del model ONDC.</p>
Exemples	Anàlisi i millora de processos

4.6.1.7.4 Atribut: *Criteri de nominalització*

Identificador	ENT007-003
Nom formal	Criteri de nominalització
Nom curt	Nominal
Domini	<p>El criteri de nominalització defineix com s'identificaran de forma unívoca els expedient dins la sèrie documental. L'assignació d'aquest criteri comporta dues conseqüències importants:</p> <ol style="list-style-type: none"> 1. El criteri de nominalització implica criteris d'ordenació dels expedients dins de la sèrie documental. De vegades això es confon amb criteris de classificació. 2. El criteri de nominalització implica també definir els criteris d'agregació per a la formació d'expedients. Això vol dir quin serà l'objecte sobre el qual s'instanciarà el procés associat a la sèrie documental. També depèn del tipus d'objecte el fet que parlem de diferents tipus de sèries documentals: reglat, tècnic i temàtic. <p>Els prefix dels criteris de nominalització podrien derivar-se en noves facetes de classificació, o també en termes d'etiquetatge en un vocabulari controlat. L'ús de la mateixa identificació d'objecte proporciona un altre punt d'accés als expedients, i és bàsic per a la normalització en el disseny global del sistema de gestió documental.</p>
Tipus de dades	Alfanumèric
Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	El nominal utilitza en el context d'aquesta tesi una notació especial:

	<ul style="list-style-type: none"> • Sempre s'inicia amb el prefix "N-", per indicar, en cas de tractaments massius de les dades, que es tracta d'un nom o número d'expedient. • El nominal pot contenir parts fixes i parts variables. Les parts fixes s'escriuen sense cap notació especial, mentre les parts variables s'escriuen entre claudàtors ("[]").
Valor per defecte	No aplica.
Fonts	<p>ISAD disposa d'un element, denominat 3.3.4 System of arrangement, que permet incloure la descripció de qualsevol forma d'organitzar el contingut d'una sèrie documental, tant si es tracta d'un criteri de classificació com d'ordenació. Malgrat el concepte de criteri de nominalització va més enllà, doncs descriu la granularitat en la formació d'expedients, en essència determina la forma d'ordenar el contingut de la sèrie documental, per la qual cosa és equivalent a aquest element d'ISAD.</p> <p>La mateixa correspondència que amb ISAD es pot trobar amb l'element d'EAD <arrangement>.</p> <p>Correspon també a l'atribut <DOC_ORGANIZA> del model NEDA.</p> <p>El concepte de criteri d'ordenació dins una sèrie documental queda recollit a RiC a l'atribut RiC-A40 Structure. Pel que fa a la construcció del criteri, els valors per a definir el criteri de nominalització en molts casos poden correspondre a l'objecte descrit, per la qual cosa es podran obtenir de l'entitat RiC-E01 Cosa. En d'altres casos el criteri de nominalització pot tenir caràcter territorial, per la qual cosa els valors per construir-lo es podrien obtenir de l'entitat RiC-E22 Lloc.</p>
Exemples	N-PRO/[Inicials projecte]

4.6.1.7.5 Atribut: Caràcter

Identificador	ENT007-004
Nom formal	Caràcter
Nom curt	Caràcter
Domini	<p>El productor d'una sèrie documental pot ser l'agent principal d'un procés (que té la "competència" per iniciar, desenvolupar o acabar el procés), o pot ser només un agent secundari que ajuda els agents principals a completar les seves tasques, però sense tenir cap competència determinant sobre el procés. En el primer cas, la sèrie documental produïda s'anomena "Primària", i en el segon cas s'anomena "Secundària".</p> <p>Tota sèrie documental primària té quatre característiques:</p> <ol style="list-style-type: none"> 1. Documenta totes les tasques essencials del procés. Una tasca és essencial quan treure aquesta tasca del procés té com a

	<p>conseqüència aturar el procés, o evita que el procés arribi al seu final normal.</p> <ol style="list-style-type: none"> 2. Totes aquestes tasques es documenten amb evidències originals, és a dir, evidències emanades directament de les tasques que s'estan executant. Aquestes evidències es poden substituir per còpies quan aquestes còpies estan autenticades i tenen la mateixa qualitat probatòria que l'evidència original. 3. L'expedient ha de ser conservat pel productor, o en un entorn on el productor pugui tenir control sobre la cadena de custòdia. 4. Totes les evidències han d'estar en un suport que permeti conservar-los enters i autèntics durant el temps que estableixi la regla de conservació. <p>El caràcter de primària o secundària d'una sèrie es refereix al caràcter intrínsec, no al que és producte d'una relació amb una altra sèrie.</p>
Tipus de dades	Alfanumèric categòric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	<p>L'atribut pot prendre dos valors:</p> <ul style="list-style-type: none"> • Primària • Secundària
Valor per defecte	"Primària"
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	Primària

4.6.1.7.6 Atribut: Descripció

Identificador	ENT007-005
Nom formal	Descripció
Nom curt	Descripció
Domini	Per a la gestió del canvi en el procés d'implantació d'un sistema de gestió documental és essencial disposar d'una descripció de què s'ha d'arxivar dins d'una sèrie documental, com identificar el seu objecte, com gestionar els sistemes d'informació relacionats amb una sèrie documental, etc.
Tipus de dades	Alfanumèric descriptiu
Mín. ocurrencies	0
Màx. ocurrencies	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	<p>Correspon a l'element ISAD 3.3.1 Scope and content.</p> <p>Correspon també a l'element EAD <scopecontent>.</p>

	<p>El model MoReq recull aquest atribut en la metadada M14.4.16 Description.</p> <p>Equival a l'atribut abstract que preveu la norma ISO 23081.</p> <p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE5 - Descripción.</p> <p>Correspon a l'atribut <DOC_SINTESIS> del model NEDA.</p> <p>RiC recull el concepte de descripció o comentaris de caràcter genèric en l'atribut RiC-A16 Descriptive note. També ho desenvolupa amb l'atribut RiC-A38 Scope and content, i per a les sèries tancades a l'atribut RiC-A21 History.</p> <p>Aquest atribut és equivalent a la metadada Description del model ONDC.</p>
Exemples	Arxivarem en aquesta sèrie tota la documentació generada pel procés de preinscripció escolar, des de la sol·licitud de preinscripció presentada per l'alumne/a fins a la formalització de la matrícula en un centre educatiu.

4.6.1.7.7 Atribut: Vitalitat

Identificador	ENT007-006
Nom formal	Vitalitat
Nom curt	Vital
Domini	<p>Es considera vital aquella sèrie documental la conservació de la qual és determinant per a la supervivència de l'organització i el manteniment de la seva activitat.</p> <p>De vegades es confon vital amb "essencial". Vital s'aplica a les sèries documentals, mentre que essencial s'aplica a les evidències.</p>
Tipus de dades	Booleà
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	L'atribut pot prendre dos valors: "Sí" i "No"
Valor per defecte	"No"
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	No

4.6.1.7.8 Atribut: 100% Digital

Identificador	ENT007-007
Nom formal	Percentatge de digitalització
Nom curt	Digital

Domini	<p>Aquest atribut recull el percentatge aproximat de documentació de la sèrie que està en format electrònic. La necessitat d'aquest atribut s'explica en el context d'un procés de digitalització de la documentació de les organitzacions, en el qual la supressió completa dels documents en paper s'ha convertit en un objectiu en si mateix per a les organitzacions.</p> <p>Es contempla aquest atribut per descriure el nivell de digitalització de la sèrie documental en aquells casos on encara no s'ha pogut descriure la composició evidencial de la sèrie i, per tant, no es pot calcular l'indicador a partir dels formats de les evidències de la sèrie.²²⁰ Aquest atribut s'ha inclòs en el model com a resultat de necessitats específiques de gestió de la informació del Consorci d'Educació de Barcelona. No es considera determinant als efectes d'anàlisi de les dades recollides.</p>
Tipus de dades	Numèric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	<p>Aquest atribut pot prendre tres valors:</p> <ul style="list-style-type: none"> • 0%, que significa que la sèrie no està en procés de transformació digital i, per tant, tota la documentació original està en suport paper. • 50%, que significa que la sèrie ha iniciat el procés de transformació digital, i part de la documentació està en format digital. • 100%, que significa que la sèrie ha completat el procés de transformació digital i ja no produeix cap document original en paper o format no digital.
Valor per defecte	50%
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	100%

4.6.1.7.9 Atribut: Data creació

Identificador	ENT007-008
Nom formal	Data de creació
Nom curt	Data creació
Domini	La data de creació d'una sèrie documental és la data en què aquesta sèrie ha estat identificada i fixada definitivament dins el mapa de sèries documentals d'una organització.
Tipus de dades	Data
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	Data vàlida.
Valor per defecte	No aplica.

²²⁰ Vegeu el capítol corresponent a l'anàlisi de les dades.

Fonts	<p>Correspon a l'element ISAD 3.1.3 Date(s), que contempla la possibilitat d'indicar les dates extremes d'una sèrie documental en un sol element.</p> <p>Correspon a l'element EAD <unitdate>.</p> <p>El model MoReq recull aquest atribut en la metadada M14.4.9 Created timestamp, tot i que també s'hi pot aplicar la metadada M14.4.61 originated Date/time.</p> <p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE4.1 - Fecha inicio.</p> <p>Correspon a l'atribut <DOC_FECHA> del model NEDA.</p> <p>Equival a l'atribut RiC-A29 Normalized date dins l'entitat RiC-E20 Date range. Vegeu nota explicativa a l'apartat Fonts de l'atribut ENT001-006.</p> <p>Aquest atribut és equivalent a la metadada Temporal coverage from del model ONDC.</p>
Exemples	12 de gener de 2008.

4.6.1.7.10 Atribut: Data supressió

Identificador	ENT007-009
Nom formal	Data de supressió
Nom curt	Data supressió
Domini	<p>La data de supressió és la data en què una sèrie documental deixa d'estar activa en el mapa de sèries d'una organització. Els motius de la supressió poden ser:</p> <ul style="list-style-type: none"> • Un canvi substancial en la naturalesa, característiques o substrat documental del procés que genera la sèrie (per exemple una nova aplicació informàtica). • Un canvi substancial en el propi procés que genera la sèrie, o la supressió d'aquest procés.
Tipus de dades	Data
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	Data vàlida.
Valor per defecte	No aplica.
Fonts	<p>Correspon a l'element ISAD 3.1.3 Date(s), que contempla la possibilitat d'indicar les dates extremes d'una sèrie documental en un sol element.</p> <p>Correspon a l'element EAD <unitdate>.</p>

	<p>El model MoReq recull aquest atribut en la metadada M14.4.5 Closed timestamp.</p> <p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE4.2 - Fecha fin.</p> <p>Correspon a l'atribut <DOC_FECHA> del model NEDA.</p> <p>Equival a l'atribut RiC-A29 Normalized date dins l'entitat RiC-E20 Date range. Vegeu nota explicativa a l'apartat Fonts de l'atribut ENT001-006.</p> <p>Aquest atribut és equivalent a la metadada Temporal coverage to del model ONDC.</p>
Exemples	31 de desembre de 2021.

4.6.1.8 Entitat: Regla d'accés

4.6.1.8.1 Descripció de l'entitat

Identificador	ENT008
Nom formal	Regla d'accés
Nom curt	Regla d'accés
Domini	<p>Entenem per regla d'accés qualsevol determinació del tipus d'accés que es pot fer sobre una sèrie documental durant un determinat període de temps, i per part d'un determinat tipus o perfil d'agent.</p> <p>Les regles d'accés s'apliquen sempre sobre agrupacions documentals completes (expedients), que corresponen a instanciacions de la sèrie documental. Puntualment es poden excepcionar determinades evidències dins l'expedient, o crear permisos i terminis que s'apliquen només sobre determinades evidències de l'expedient.</p>
Atributs	<p>ENT008-001 Identificador</p> <p>ENT008-002 Esdeveniment d'inici o disparador</p> <p>ENT008-003 Termini</p> <p>ENT008-004 Resolució</p> <p>ENT008-005 Motiu de limitació</p> <p>ENT008-006 Criteri d'aplicació de la regla</p> <p>ENT008-007 Criteri de publicitat per transparència</p> <p>ENT008-008 Data creació</p> <p>ENT008-009 Data supressió</p>
Relacions	<p>REL005 Atorga permisos d'accés a un o més agents</p> <p>REL014 Regula l'accés a una sèrie documental</p>
Fonts	<p>La norma ISAD disposa d'una àrea completa per a la descripció de les condicions d'accés i ús. En relació amb el nostre model, l'equivalència es concentra en dos elements d'ISAD: l'element 3.4.1 Conditions governing access, i l'element 3.4.2 Conditions governing reproduction.</p> <p>EAD disposa de dos elements, <accessrestrict> i <userrestrict>, que es poden considerar equivalents també a aquesta entitat.</p> <p>MoReq contempla l'existència de la regla d'accés des de la seva perspectiva tecnològica dins un gestor documental (<i>Access Control List</i>, o <i>ACL</i>). Ho fa a través de dos elements que no categoritza com a entitats sinó com a estructures de dades (<i>data structures</i>), i que són D14.3.1 Access Control Entry, i també D14.3.2 Access Control list.</p> <p>Es pot posar en relació aquesta entitat amb la propietat 11 Permissions del model AGRkMS.</p> <p>La norma ISO 23081 contempla Mandates com una de les cinc entitats principals del seu model, en els qual inclou les regles que determinen el cicle de vida.</p>

	<p>Aquesta entitat es pot posar en relació amb la metadada eEMGDE8 – Seguridad del model e-EMGDE, però també es pot relacionar amb la metadada eEMGDE9 - Derechos de acceso, uso y reutilización.</p> <p>Correspon a l'atribut <DOC_ACCESO> del model NEDA.</p> <p>L'entitat en conjunt és expressada a RiC mitjançant dos atributs, RiC-A08 Conditions of acces, i RiC-A09 Conditions of use. Aquesta equivalència és extrapolable a diferents atributs d'aquesta entitat.</p> <p>La metadada Access Rights del model ONDC es pot considerar també equivalent a aquesta entitat.</p>
--	---

4.6.1.8.2 Atribut: Identificador

Identificador	ENT008-001
Nom formal	Identificador
Nom curt	Id
Domini	Identificador únic de la regla d'accés. Agafa com a base el codi de la sèrie documental a la qual determina.
Tipus de dades	Alfanumèric
Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	<p>Es tracta d'un codi qualificat, que es compona seguint el següent patró:</p> <ul style="list-style-type: none"> • Codi complet de la sèrie documental • Guió • Número seqüencial que identifica la regla dins la seva sèrie documental (tres dígits)
Valor per defecte	No aplica.
Fonts	<p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE8.1.2 - Código de la Política de control de acceso.</p> <p>Correspon a l'atribut genèric RiC-A22 Identifier.</p>
Exemples	ASS005-001-001

4.6.1.8.3 Atribut: Esdeveniment d'inici o disparador

Identificador	ENT008-002
Nom formal	Esdeveniment d'inici o disparador
Nom curt	Disparador
Domini	Descripció de l'esdeveniment que activa la regla, és a dir que inicia l'aplicació de la regla sobre un conjunt documental determinat. Pot ser l'inici de la tramitació, un canvi en un estat d'un expedient, etc.
Tipus de dades	Alfanumèric
Mín. ocurrències	1

Màx. ocurrencies	10
Regles de validació	No
Valor per defecte	No aplica.
Fonts	Es pot considerar equivalent a l'atribut trigger que inclou la norma ISO 23081 . L'eventual categorització d'aquest atribut el faria equiparable a l'atribut RiC-A18 Event type .
Exemples	Tancament d'expedient

4.6.1.8.4 Atribut: *Termini (o període)*

Identificador	ENT008-003
Nom formal	Termini
Nom curt	Termini
Domini	Termini indicat en una unitat de mesura de temps a comptar des del valor de l'atribut ENT008-002 per a l'aplicació de la disposició descrita en l'atribut ENT008-004.
Tipus de dades	Alfanumèric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	El termini d'aplicació de la regla es pot determinar per calendari (termini precalculat) o per un esdeveniment (termini detectat). En cas que es determini per calendari, s'indicarà en aquest atribut el termini, habitualment calculat en anys, a comptar a partir de la data en què s'ha esdevingut l'esdeveniment d'inici. En cas que es determini per esdeveniment, es farà constar en aquest atribut la descripció de l'esdeveniment que posa fi al termini.
Valor per defecte	No aplica.
Fonts	Equivalent a RiC-A08 Conditions of acces , i RiC-A09 Conditions of use .
Exemples	3 mesos

4.6.1.8.5 Atribut: *Resolució (o disposició)*

Identificador	ENT008-004
Nom formal	Resolució
Nom curt	Resolució
Domini	Forma en què un determinat agent accedeix als documents d'aquella sèrie documental durant el termini definit, i amb quins permisos o capacitats ho fa.
Tipus de dades	Alfanumèric categòric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	Llista amb els següents valors:

	<ul style="list-style-type: none"> • Esborrar • Crear • Modificar • Publicar • Versionar • Visualitzar • Visualitzar metadades completes • Visualitzar descripció • Sense accés <p>Els permisos són incrementals, de major a menor permís.</p>
Valor per defecte	No aplica.
Fonts	<p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE8.1.1 - Nivel de acceso. També es pot posar en relació amb la metadada eEMGDE8.3 – Permisos, i també amb eEMGDE9.1 - Tipo de acceso.</p> <p>Equivalent a RiC-A08 Conditions of acces, i RiC-A09 Conditions of use.</p>
Exemples	Visualitzar descripció

4.6.1.8.6 Atribut: Motiu de limitació

Identificador	ENT008-005
Nom formal	Motiu de limitació
Nom curt	Limitació
Domini	Descripció del motiu que justifica no haver aplicat un permís mínim de visualització o superior.
Tipus de dades	Alfanumèric categòric
Mín. ocurrencies	0
Màx. ocurrencies	1
Regles de validació	<p>Llista de valors. A dia d'avui els valors entre els quals es pot triar són els següents:</p> <ul style="list-style-type: none"> • Protecció de dades personals • Propietat intel·lectual • Segregació escolar
Valor per defecte	No aplica.
Fonts	<p>El model e-EMGDE contempla per aquest atribut dues metadades: eEMGDE9.2 - Código de la causa de limitación, i també eEMGDE9.3 - Causa legal/normativa de limitación.</p> <p>Equivalent a RiC-A08 Conditions of acces, i RiC-A09 Conditions of use.</p>
Exemples	Segregació escolar

4.6.1.8.7 Atribut: Criteri d'aplicació de la regla

Identificador	ENT008-006
Nom formal	Criteri d'aplicació de la regla
Nom curt	Criteri d'aplicació
Domini	Detalls textuals relatius a la forma com cal aplicar la regla, que poden incloure incidències i altres aspectes a destacar. S'inclou en aquest atribut l'esment a excepcions puntuals en els documents o evidències sobre els quals s'aplica la regla.
Tipus de dades	Alfanumèric descriptiu
Mín. ocurrences	0
Màx. ocurrences	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	Equivalent a RiC-A08 Conditions of acces , i RiC-A09 Conditions of use .
Exemples	Accés restringit, sense perjudici d'accés parcial. Motivació: conté dades de caràcter personal especialment protegides.

4.6.1.8.8 Atribut: Criteri de publicitat per transparència

Identificador	ENT008-007
Nom formal	Criteri de publicitat per transparència
Nom curt	Criteri de publicitat
Domini	Descripció dels requisits pel que fa a l'aplicació de la normativa de transparència i necessitat de publicar de forma proactiva informació continguda en aquella sèrie documental.
Tipus de dades	Alfanumèric descriptiu
Mín. ocurrences	0
Màx. ocurrences	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	El model e-EMGDE contempla per aquest atribut la metadada eEMGDE8.1.1 - Nivel de acceso . I també amb la metadada eEMGDE9.4 - Condiciones de reutilización . Equivalent a RiC-A08 Conditions of acces , i RiC-A09 Conditions of use .
Exemples	Publicar en format de dades obertes.

4.6.1.8.9 Atribut: Data creació

Identificador	ENT008-008
Nom formal	Data de creació
Nom curt	Data creació
Domini	Data a partir de la qual es considera que la regla està activa i s'aplica sobre la sèrie documental.
Tipus de dades	Data
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	Data vàlida.
Valor per defecte	No aplica.
Fonts	Equival a l'atribut RiC-A29 Normalized date dins l'entitat RiC-E20 Date range . Vegeu nota explicativa a l'apartat Fonts de l'atribut ENT001-006 .
Exemples	12 de gener de 2008.

4.6.1.8.10 Atribut: Data supressió

Identificador	ENT008-009
Nom formal	Data de supressió
Nom curt	Data supressió
Domini	Data a partir de la qual es considera que la regla ja no està activa i ja no s'aplica sobre la sèrie documental.
Tipus de dades	Data
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	Data vàlida.
Valor per defecte	No aplica.
Fonts	Equival a l'atribut RiC-A29 Normalized date dins l'entitat RiC-E20 Date range . Vegeu nota explicativa a l'apartat Fonts de l'atribut ENT001-006 .
Exemples	31 de desembre de 2021.

4.6.1.9 Entitat: Regla de conservació

4.6.1.9.1 Descripció de l'entitat

Identificador	ENT009
Nom formal	Regla de conservació
Nom curt	Regla de conservació
Domini	Entenem per regla de conservació qualsevol regla que determina les condicions de conservació d'una sèrie documental durant un termini. Aquestes condicions poden implicar el lloc, el format o la supervivència de la sèrie (en el cas de regles d'eliminació). Quan les condicions de conservació són diferents per a diferents evidències de la sèrie documental durant un mateix termini, es crearan tantes regles per aquell termini com condicions de conservació calgui determinar.
Atributs	<p>ENT009-001 Identificador</p> <p>ENT009-002 Esdeveniment d'inici o disparador</p> <p>ENT009-003 Termini</p> <p>ENT009-004 Resolució</p> <p>ENT009-005 Criteri de mostreig</p> <p>ENT009-006 Suport de destí</p> <p>ENT009-007 Format de destí</p> <p>ENT009-008 Lloc de destí</p> <p>ENT009-009 Data creació</p> <p>ENT009-010 Data supressió</p>
Relacions	REL015 Regula la conservació d'una sèrie documental
Fonts	<p>La informació descriptiva continguda en aquesta entitat correspon al que la norma ISAD descriu en l'element 3.3.2 Appraisal, destruction and scheduling information.</p> <p>EAD disposa d'un element equiparable a aquesta entitat, l'element <appraisal>.</p> <p>MoReq inclou el concepte de regla de conservació amb l'entitat E14.2.6 Disposal Schedule. MoReq detalla especialment aquesta entitat amb un nombre considerable de metadades previstes. Es detallen a continuació aquests vincles.</p> <p>Aquesta entitat es pot posar en relació amb la propietat 18 Disposal del model AGRkMS.</p> <p>La norma ISO 23081 contempla Mandates com una de les cinc entitats principals del seu model, en els qual inclou les regles que determinen el cicle de vida.</p> <p>El model e-EMGDE disposa d'una equivalència per aquesta entitat en la metadada eEMGDE13 – Calificación.</p> <p>Correspon a l'atribut <DOC_VALORA> del model NEDA.</p>

	La metadada Disposal del model ONDC es pot considerar també equivalent a aquesta entitat.
--	---

4.6.1.9.2 Atribut: *Identificador*

Identificador	ENT009-001
Nom formal	Identificador
Nom curt	Id
Domini	Identificador únic de la regla de conservació. Agafa com a base el codi de la sèrie documental a la qual determina.
Tipus de dades	Alfanumèric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	Es tracta d'un codi qualificat, que es compon seguint el següent patró: <ul style="list-style-type: none"> • Codi complet de la sèrie documental • Guió • Número seqüencial que identifica la regla dins la seva sèrie documental (tres dígits)
Valor per defecte	No aplica.
Fonts	El model MoReq recull aquest atribut en la metadada M14.4.100 System Identifier . Correspon a l'atribut genèric RiC-A22 Identifier .
Exemples	INF003-001-001

4.6.1.9.3 Atribut: *Esdeveniment d'inici o disparador*

Identificador	ENT009-002
Nom formal	Esdeveniment d'inici o disparador
Nom curt	Disparador
Domini	Descripció de l'esdeveniment que activa la regla, és a dir que inicia l'aplicació de la regla sobre un conjunt documental determinat. Pot ser l'inici de la tramitació, un canvi en un estat d'un expedient, etc.
Tipus de dades	Alfanumèric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	Es pot considerar equivalent a l'atribut trigger que inclou la norma ISO 23081 .

	<p>MoReq contempla diverses metadades que permeten descriure aspectes de l'esdeveniment d'inici i disparador de la regla. Les que tenen una relació més directa amb aquest atribut són M14.4.95 Retention trigger Element Identifier, M14.4.94 Retention trigger Code, M14.4.93 Retention Start Date i M14.4.18 Disposal Action Code.</p> <p>L'eventual categorització d'aquest atribut el faria equiparable a l'atribut RiC-A18 Event type.</p>
Exemples	Tancament d'expedient

4.6.1.9.4 Atribut: *Termini (o període)*

Identificador	ENT009-003
Nom formal	Termini
Nom curt	Termini
Domini	Termini indicat en una unitat de mesura de temps a comptar des del valor de l'atribut ENT009-002 per a l'aplicació de la disposició descrita en l'atribut ENT009-004.
Tipus de dades	Alfanumèric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	El termini d'aplicació de la regla es pot determinar per calendari (termini precalculat) o per un esdeveniment (termini detectat). En cas que es determini per calendari, s'indicarà en aquest atribut el termini, habitualment calculat en anys, a comptar a partir de la data en què s'ha esdevingut l'esdeveniment d'inici. En cas que es determini per esdeveniment, es farà constar en aquest atribut la descripció de l'esdeveniment que posa fi al termini.
Valor per defecte	No aplica.
Fonts	El model e-EMGDE contempla per aquest atribut la metadada eEMGDE13.1.1.2 – Plazo , com també les metadades eEMGDE13.2.3 - Plazo de ejecución de la acción dictaminada i eEMGDE13.3.2 - Plazo de transferencia .
Exemples	2 anys

4.6.1.9.5 Atribut: *Resolució (o disposició)*

Identificador	ENT009-004
Nom formal	Resolució
Nom curt	Resolució
Domini	Determinació de la forma d'existència de les evidències durant el termini indicat. Aquesta forma pot implicar la no-existència (regla d'eliminació) o l'existència substancialment diferent (regla de transformació).
Tipus de dades	Alfanumèric categòric

Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	Llista amb els següents valors: <ul style="list-style-type: none"> • Transferir (moure o canviar de lloc d'emmagatzemament) • Transformar o convertir (canviar de suport i/o format digital) • Eliminar
Valor per defecte	No aplica.
Fonts	MoReq contempla diverses metadades per indicar el termini de conservació. Les que tenen una relació més directa són M14.4.89 Retention Period Duration Number, M14.4.90 Retention Period Interval Code, M14.4.19 Disposal Action Due Date, i M14.4.20 Disposal Confirmation Due Date. El model e-EMGDE contempla per aquest atribut la metadada eEMGDE13.2.2 - Acción dictaminada.
Exemples	Eliminar

4.6.1.9.6 Atribut: Criteri de mostreig

Identificador	ENT009-005
Nom formal	Criteri de mostreig
Nom curt	Mostreig
Domini	Descripció del conjunt d'evidències (dins de l'expedient) o del conjunt d'expedients (dins la sèrie documental) que seran objecte de la resolució indicada en l'atribut ENT009-004. Aquesta tria pot respondre a un model: <ul style="list-style-type: none"> • Exemplar/selectiu (mostra obtinguda per selecció manual amb un criteri personal no sistemàtic) • Sistemàtic (mostra precalculada en funció d'un criteri preestablert a partir del criteri de nominalització o d'una tipologia d'evidències) • Aleatori (mostra estadísticament representativa obtinguda a partir d'un determinat criteri de nominalització o d'una tipologia d'evidències) <p>Tot mostreig donarà origen a més d'una regla de conservació per un mateix període, i per tantes mostres com aquelles en què s'hagi dividit el conjunt de la població documental/evidencial.</p>
Tipus de dades	Alfanumèric descriptiu
Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.

Exemples	Mostra: queixes presentades per clients el primer cognom dels quals comença per A o per R.
----------	--

4.6.1.9.7 Atribut: Suport de destí

Identificador	ENT009-006
Nom formal	Suport de destí
Nom curt	Suport
Domini	Suport físic sobre el qual estaran els documents/evidències durant aquest termini. Aquest atribut només s'informarà quan la resolució sigui transformar o convertir a un suport diferent al suport original de les evidències. Es pot aplicar també sobre una mostra (tots els documents que estan en un determinat suport).
Tipus de dades	Alfanumèric categòric
Mín. ocurrencies	0
Màx. ocurrencies	1
Regles de validació	Aquest atribut pot prendre dos valors: <ul style="list-style-type: none"> • Paper • Digital
Valor per defecte	No aplica.
Fonts	
Exemples	Digital

4.6.1.9.8 Atribut: Format de destí

Identificador	ENT009-007
Nom formal	Format de destí
Nom curt	Format
Domini	Format digital en el qual estaran els documents/evidències durant aquest termini. Aquest atribut només s'informarà quan la resolució sigui transformar o convertir a un format diferent al suport original de les evidències. Es pot aplicar també sobre una mostra (tots els documents que estan en un determinat format).
Tipus de dades	Alfanumèric categòric
Mín. ocurrencies	0
Màx. ocurrencies	1
Regles de validació	Aquest atribut només s'informa quan la resolució implica conservació. L'atribut es valida contra una llista de valors creada a partir dels formats utilitzats per l'organització. En aquesta llista es contempen les següents casuístiques: <ul style="list-style-type: none"> • Formularis o models de document en suports no digitals: s'informa el suport. • Formularis o models de document estàtics o en forma de document-objecte: s'informa el format de fitxer amb el nom i

	<p>identificador del registre de formats PRONOM (<i>PRONOM Unique Identifier</i>, o PUID).</p> <ul style="list-style-type: none"> • Formularis interactius o vinculats a aplicacions informàtiques: s'informa el nom de la tecnologia o plataforma que s'ha utilitzat per a desenvolupar el formulari.
Valor per defecte	No aplica.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	Portable Document Format (PDF)

4.6.1.9.9 Atribut: Lloc de destí

Identificador	ENT009-008
Nom formal	Lloc de destí
Nom curt	Lloc
Domini	Indret físic o repositori digital el qual estaran els documents/evidències durant aquest termini. Aquest atribut només s'informarà quan la resolució sigui transferir. Es pot aplicar també sobre una mostra.
Tipus de dades	Alfanumèric categòric
Mín. ocurrencies	0
Màx. ocurrencies	1
Regles de validació	Només s'informa quan la resolució implica conservació. Es valida contra la llista de tots els repositoris digitals o dipòsits físics que l'organització utilitzi per a l'arxiu de les seves evidències, tant siguin propis com externs.
Valor per defecte	No aplica.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	Arxiu electrònic únic

4.6.1.9.10 Atribut: Data creació

Identificador	ENT009-009
Nom formal	Data de creació
Nom curt	Data creació
Domini	Data a partir de la qual es considera que la regla està activa i s'aplica sobre la sèrie documental.
Tipus de dades	Data
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	Data vàlida.
Valor per defecte	No aplica.
Fonts	El model MoReq recull aquest atribut en la metadada M14.4.9 Created timestamp , tot i que també s'hi pot aplicar la metadada M14.4.61 originated Date/time .

	Equival a l'atribut RiC-A29 Normalized date dins l'entitat RiC-E20 Date range . Vegeu nota explicativa a l'apartat Fonts de l'atribut ENT001-006 .
Exemples	12 de gener de 2008.

4.6.1.9.11 Atribut: Data supressió

Identificador	ENT009-010
Nom formal	Data de supressió
Nom curt	Data supressió
Domini	Data a partir de la qual es considera que la regla ja no està activa i ja no s'aplica sobre la sèrie documental.
Tipus de dades	Data
Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	Data vàlida.
Valor per defecte	No aplica.
Fonts	El model MoReq recull aquest atribut en la metadada M14.4.5 Closed timestamp . Equival a l'atribut RiC-A29 Normalized date dins l'entitat RiC-E20 Date range . Vegeu nota explicativa a l'apartat Fonts de l'atribut ENT001-006 .
Exemples	31 de desembre de 2021.

4.6.1.10 Entitat: Disposició normativa

4.6.1.10.1 Descripció de l'entitat

Identificador	ENT010
Nom formal	Disposició normativa
Nom curt	Disposició
Domini	Regulació normativa o mandat que determina en algun aspecte les característiques d'altres elements del sistema de governança, a través del procés amb el qual estan vinculats.
Atributs	ENT010-001 Identificador ENT010-002 Títol ENT010-003 Rang ENT010-004 Àmbit ENT010-005 Data vigència ENT010-006 Data derogació
Relacions	REL010 Pot regular un o més processos
Fonts	<p>El model AGRkMS contempla l'entitat Mandate, que defineix com “<i>A source of business requirements, including recordkeeping requirements</i>”.</p> <p>La norma ISO 23081 contempla Mandates com una de les cinc entitats principals del seu model.</p> <p>El model e-EMGDE també contempla una entitat equivalent a aquesta, que anomena Regulación.</p> <p>El model NEDA contempla aquesta entitat amb l'entitat Norma.</p> <p>RiC contempla l'entitat RiC-E16 Rule, la qual es desplega amb l'entitat RiC-E17 Mandate. Ambdues defineixen qualsevol tipus de norma de caràcter escrit o consuetudinari que reguli l'ús de la informació.</p> <p>La metadada Legal authority del model ONDC es pot considerar també equivalent a aquesta entitat.</p>

4.6.1.10.2 Atribut: Identificador

Identificador	ENT010-001
Nom formal	Identificador
Nom curt	Id
Domini	Identificador assignat a cada disposició o mandat en el diari oficial o document jurídic corresponent. En cas que no existeixi aquesta nomenclatura, se substituirà per un codi compost localment.
Tipus de dades	Alfanumèric

Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	<p>La pauta per construir l'identificador és la següent:</p> <p>En cas que la disposició disposi de codi propi, l'identificador es compondrà seguint el següent patró:</p> <ul style="list-style-type: none"> • Inicial del diari oficial (quatre caràcters) • Guió • Tipus de disposició (tres primers caràcters) • Guió • Número de la disposició al diari oficial corresponent <p>En cas que la disposició no disposi de codi propi, l'identificador es compondrà seguint el següent patró:</p> <ul style="list-style-type: none"> • Inicial de Administració que l'expedeix • Guió • Inicials de l'òrgan que l'expedeix • Guió • Tipus de disposició (tres primers caràcters) • Guió • Data de la disposició en format any-mes-dia (sense separadors)
Valor per defecte	No aplica.
Fonts	<p>Aquest atribut equival a la propietat 2 Identifier del model AGRkMS.</p> <p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE2 – Identificador.</p> <p>Equival a l'atribut identifier que preveu la norma ISO 23081.</p> <p>Correspon a l'atribut genèric RiC-A22 Identifier.</p>
Exemples	DEC_102/2010

4.6.1.10.3 Atribut: Títol

Identificador	ENT010-002
Nom formal	Títol
Nom curt	Títol
Domini	Títol formal complet atribuït a la disposició.
Tipus de dades	Alfanumèric
Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	<p>Aquest atribut equival a la propietat 3 Name del model AGRkMS.</p> <p>Equival a l'atribut title que preveu la norma ISO 23081.</p>

	<p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE3 – Nombre.</p> <p>Correspon a l'atribut genèric RiC-A28 Name.</p>
Exemples	Llei 22/1998, de 30 de desembre, de la Carta Municipal de Barcelona.

4.6.1.10.4 Atribut: Rang

Identificador	ENT010-003
Nom formal	Rang
Nom curt	Rang
Domini	Rang de la norma o disposició, en funció de la jurisdicció on apliqui.
Tipus de dades	Alfanumèric categòric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	<p>L'atribut pot contenir un dels següents valors:</p> <ul style="list-style-type: none"> • Llei • Decret • Ordre • Disposició no publicada
Valor per defecte	No aplica.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	Llei

4.6.1.10.5 Atribut: Àmbit

Identificador	ENT010-004
Nom formal	Àmbit
Nom curt	Àmbit
Domini	Àmbit territorial sobre el qual aplica la disposició.
Tipus de dades	Alfanumèric categòric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	<p>L'atribut pot contenir un dels següents valors:</p> <ul style="list-style-type: none"> • Local • Autonòmic • Estatal • Europeu • Internacional
Valor per defecte	No aplica.

Fonts	<p>Aquest atribut equival a la propietat 8 Jurisdiction del model AGRkMS, tot i que també es pot posar en relació amb la propietat 16 Coverage.</p> <p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE7 - Àmbito.</p> <p>Equival a l'atribut jurisdiction que preveu la norma ISO 23081.</p> <p>Es pot equiparar a l'atribut RiC-A32 Place type.</p>
Exemples	Local

4.6.1.10.6 Atribut: Data vigència

Identificador	ENT010-005
Nom formal	Data de vigència
Nom curt	Data de vigència
Domini	Data de publicació de la disposició en el diari oficial corresponent. En el cas de les disposicions no publicades, data de signatura de la disposició.
Tipus de dades	Data
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	Data vàlida.
Valor per defecte	No aplica.
Fonts	<p>Aquest atribut equival a la propietat 4 Date range del model AGRkMS.</p> <p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE4.1 - Fecha inicio.</p> <p>Equival a l'atribut RiC-A29 Normalized date dins l'entitat RiC-E20 Date range. Vegeu nota explicativa a l'apartat Fonts de l'atribut ENT001-006.</p>
Exemples	12 de gener de 2008.

4.6.1.10.7 Atribut: Data derogació

Identificador	ENT010-006
Nom formal	Data de derogació
Nom curt	Data de derogació
Domini	Data de publicació en el diari oficial corresponent de la disposició que deroga aquesta disposició. En el cas de les disposicions no publicades, data de signatura de la disposició que deroga aquesta disposició. En els casos de disposicions que inclouen un termini de vigència preestablert, data en què la pròpia disposició determina la fi del termini de vigència.

Tipus de dades	Data
Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	Data vàlida.
Valor per defecte	No aplica.
Fonts	<p>Aquest atribut equival a la propietat 4 Date range del model AGRkMS.</p> <p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE4.2 - Fecha fin.</p> <p>Equival a l'atribut RiC-A29 Normalized date dins l'entitat RiC-E20 Date range. Vegeu nota explicativa a l'apartat Fonts de l'atribut ENT001-006.</p>
Exemples	31 de desembre de 2021.

4.6.1.11 Entitat: Sistema d'informació o aplicació

4.6.1.11.1 Descripció de l'entitat

Identificador	ENT011
Nom formal	Sistema d'informació o aplicació
Nom curt	Sistema d'informació
Domini	<p>El concepte de sistema d'informació és un concepte extremadament ampli, i pot contemplar des de la visió més àmplia i neutral (un conjunt d'elements, que pot incloure persones, processos, instruments, etc., que gestiona informació en una organització) fins a una infraestructura tecnològica concreta. En el nostre cas, acotem el concepte als següents límits:</p> <ul style="list-style-type: none"> • Es tracta d'una eina informàtica, que es pot equiparar a una aplicació o a qualsevol infraestructura de base de dades • Que emmagatzema informació, sigui en format estructurat (dades) o en format no estructurat (documents-objecte) • Que té uns límits clarament o convencionalment identificables <p>Quan dins una mateixa aplicació o plataforma es poden identificar diferents sistemes, habitualment s'aplicarà com a criteri d'identificació per a la segmentació els diferents processos pels quals s'utilitza cada sistema.</p>
Atributs	<p>ENT011-001 Identificador ENT011-002 Nom ENT011-003 Nom curt ENT011-004 Responsable funcional ENT011-005 Gestor TIC ENT011-006 Proveïdor ENT011-007 Entorn ENT011-008 Entorn tecnològic ENT011-009 Model de servei ENT011-010 Abast ENT011-011 Restriccions semàntiques ENT011-012 Restriccions de format ENT011-013 Controls d'integritat ENT011-014 Controls d'accés ENT011-015 Classificació de seguretat ENT011-016 Descripció ENT011-017 Funcionalitat ENT011-018 Ubicació ENT011-019 Data inici servei ENT011-020 Data fi servei</p>
Relacions	<p>REL006 És promogut per un agent REL019 Pot contenir una o més evidències ENT012 Es pot comunicar amb un o més sistemes d'informació a través d'una o més interfícies REL023 Conté una o més dades elementals REL024 Descriu de forma principal un objecte</p>

<p>Fonts</p>	<p>El model AGRkMS contempla, dins l'entitat Agent, un tipus específic d'agent que anomena System, en referència a sistema d'informació.</p> <p>Aquesta assimilació també apareix en el model e-EMGDE, el qual dins l'entitat Agente identifica un possible subtipus, que anomena Dispositivo, i que defineix com "<i>Mecanismo físico, sistema electrónico o sistema de información que lleva a cabo actuaciones de gestión</i>".</p> <p>RiC contempla una entitat específica per a descriure qualsevol element tecnològic que pugui actuar sobre la informació: RiC-E13 Mechanism. Aquesta entitat és clarament equiparable a la de sistema d'informació o aplicació informàtica. El model RiC indica a més que els <i>mecanismos</i> executen les seves activitats sempre en base a unes regles definides per agents humans. La diferència amb el nostre model rau en el fet que RiC posa el focus en el caràcter actiu del mecanisme, atès que el fa dependre de l'entitat RiC-E07 Agent, mentre que en el nostre cas el sistema d'informació o aplicació té caràcter passiu, i el recollim només en aquells casos en què conté informació, com a element bàsic per a determinar el llinatge de les dades.</p> <p>El nostre concepte de sistema d'informació requereix que el sistema tingui capacitat d'emmagatzemar dades. Qualsevol magatzem de dades serà considerat doncs, als efectes del nostre mapatge, un sistema d'informació. Això implica també que qualsevol base de dades serà considerada un sistema d'informació. Ara bé, definir els límits de la base de dades dependrà del model tecnològic en el qual estigui implementada. Per a les bases de dades relacionals, determinarem els límits utilitzant el plantejament del principi G_3.2-1 del model SIARD (les bases de dades com a documents). En el nostre model hem incorporat metadades SIARD preferentment a nivell de base de dades (requeriment M_5.1-1).</p> <p>El model més habitual de catàleg de dades contempla tanmateix només formes estàtiques de magatzems de dades. De fet, DCAT té com a principal objecte de descripció el conjunt de dades (<i>dataset</i>), juntament amb els serveis de dades. DCAT distingeix entre una versió abstracta del dataset, i les seves instanciacions o representacions tecnològiques (que anomena <i>distributions</i>), que poden ser en forma de full de càlcul, de document XML, de taula en HTML incrustada en una pàgina Web, etc. En el nostre model, el concepte de <i>distribution</i> d'un <i>dataset</i>, representat amb la classe dcat:Distribution, correspon a l'entitat sistema d'informació.</p> <p>El model ONDC, malgrat estar orientat a descriure elements estàtics (actius d'informació, i especialment "<i>datasets</i>") també contempla la possibilitat de descriure aplicacions informàtiques o equivalents. En tot cas, l'assimilació amb conjunt de dades original és clara.</p>
--------------	---

4.6.1.11.2 Atribut: Identificador (codi)

Identificador	ENT011-001
Nom formal	Identificador
Nom curt	ID
Domini	Codi que identifica de forma unívoca el sistema dins de l'organització.
Tipus de dades	Alfanumèric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	<p>El codi del sistema d'informació és un codi qualificat compostat dels següents elements:</p> <ul style="list-style-type: none">• Inicials "SI"• Tipus de sistema, indicat per dos dígits numèrics, d'acord amb el seu entorn (vegeu l'equivalència de valors en l'atribut ENT011-007)• Guió• Número seqüencial de quatre dígits
Valor per defecte	No aplica.
Fonts	<p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE2 – Identificador.</p> <p>Correspon a l'atribut genèric RiC-A22 Identifier.</p> <p>DCAT proporciona l'element de Dublin Core dct:identifier per assignar un identificador únic a cada recurs d'un catàleg.</p> <p>Aquest atribut és equivalent a la metadada Identifier del model ONDC.</p>
Exemples	SI02-0142

4.6.1.11.3 Atribut: Nom

Identificador	ENT011-002
Nom formal	Nom
Nom curt	Nom
Domini	Nom complet del sistema. Habitualment és el nom amb el que és conegut dins l'organització pels entorns o unitats que han participat en la seva concepció i desenvolupament.
Tipus de dades	Alfanumèric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	Aquest atribut equival a la propietat 3 Name del model AGRkMS .

	<p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE3 – Nombre.</p> <p>Correspon a l'atribut genèric RiC-A28 Name.</p> <p>En el cas de bases de dades és equiparable a la metadada M_5.1-1 dbname de SIARD.</p> <p>DCAT proposa l'ús de l'element de Dublin Core dct:title per indicar el nom del recurs que conté les dades.</p> <p>Aquest atribut és equivalent a la metadada Title del model ONDC.</p>
Exemples	Sistema de gestió de la preinscripció escolar en centres sufragats amb fons públics (GEDAC)

4.6.1.11.4 Atribut: Nom curt

Identificador	ENT011-003
Nom formal	Denominació d'ús habitual
Nom curt	Nom curt
Domini	Nom curt o inicials amb les quals es coneix el sistema dins l'organització. Habitualment correspon amb el nom que li donen els usuaris al sistema.
Tipus de dades	Alfanumèric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	GEDAC

4.6.1.11.5 Atribut: Responsable funcional

Identificador	ENT011-004
Nom formal	Responsable funcional
Nom curt	Responsable funcional
Domini	<p>Nom i cognoms de la persona que dins l'organització es responsabilitza de la gestió i administració funcional del sistema. Habitualment és una persona de l'àrea de negoci, que pot ocupar-se del manteniment de diccionaris, alta, baixa i modificació d'usuaris, i seguiment d'incidències. No es considera un tipus de relació entre agent i sistema, atès que els agents sempre corresponen a identitats col·lectives, o bé a rols, i no a individus concrets.</p> <p>Aquest atribut s'ha inclòs en el model com a resultat de necessitats específiques de gestió de la informació del Consorci d'Educació de</p>

	Barcelona. No es considera determinant als efectes d'anàlisi de les dades recollides.
Tipus de dades	Usuari
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	DCAT preveu un element anomenat dc:contactPoint que permet indicar una persona de referència per al recurs.
Exemples	Jordi Serra

4.6.1.11.6 Atribut: Gestor TIC

Identificador	ENT011-005
Nom formal	Responsable tecnològic
Nom curt	Gestor TIC
Domini	<p>Nom i cognoms de la persona que dins l'organització es responsabilitza del manteniment tecnològic del sistema. Habitualment és una persona de l'àrea de tecnologies, que pot ocupar-se del monitoratge del rendiment, de la resolució d'incidències, del manteniment de les taules mestres, de l'actualització de la plataforma i de la gestió dels evolutius. No es considera un tipus de relació entre agent i sistema, atès que els agents sempre corresponen a identitats col·lectives, o bé a rols, i no a individus concrets.</p> <p>Aquest atribut s'ha inclòs en el model com a resultat de necessitats específiques de gestió de la informació del Consorci d'Educació de Barcelona. No es considera determinant als efectes d'anàlisi de les dades recollides.</p>
Tipus de dades	Usuari
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	Josep Pou

4.6.1.11.7 Atribut: Proveïdor

Identificador	ENT011-006
Nom formal	Proveïdor
Nom curt	Proveïdor
Domini	Nom de la persona o empresa que s'ocupa del desenvolupament del sistema, de la correcció dels errors causants d'incidències, i del desenvolupament dels evolutius.

Tipus de dades	Alfanumèric categòric
Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	El valor d'aquest atribut es valida contra un diccionari de proveïdors autoritzats.
Valor per defecte	No aplica.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	TICxCAT

4.6.1.11.8 Atribut: Entorn

Identificador	ENT011-007
Nom formal	Entorn
Nom curt	Entorn
Domini	L'entorn és l'atribut principal per a la categorització dels sistemes d'informació. El terme entorn és utilitzat perquè condensa els dos elements que inclou aquesta categoria: <ul style="list-style-type: none"> • El tipus de tecnologia en què es basa el sistema • L'àmbit funcional en el qual es desenvolupa el sistema
Tipus de dades	Alfanumèric categòric
Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	L'atribut pot prendre tres valors: <ul style="list-style-type: none"> • Eina ofimàtica • Sistema d'informació/aplicació • Plataforma
Valor per defecte	No aplica.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	Eina ofimàtica

4.6.1.11.9 Atribut: Entorn tecnològic

Identificador	ENT011-008
Nom formal	Entorn tecnològic
Nom curt	Tecnologia
Domini	Aquest atribut descriu la o les tecnologies en què està desenvolupat el sistema d'informació que és objecte de descripció.
Tipus de dades	Alfanumèric categòric
Mín. ocurrències	1
Màx. ocurrències	1

Regles de validació	Aquest atribut s'alimenta d'una llista de valors que conté les diferents tecnologies que es vagin utilitzant a les aplicacions de l'organització.
Valor per defecte	No aplica.
Fonts	<p>RiC disposa d'un atribut que permet descriure la tecnologia d'un sistema d'informació, l'atribut RiC-A41 Technical characteristics.</p> <p>En el cas de bases de dades relacionals, aquest atribut pot ser equiparable a la metadada M_5.1-1 databaseProduct de SIARD.</p> <p>DCAT proposa l'element de Dublin Core dct:format per descriure el format i característiques físiques del recurs. Però també contempla l'element dcat:distribution per a representar qualsevol implementació tecnològica d'un <i>dataset</i>.</p> <p>Aquest atribut és equivalent a la metadada Format del model ONDC.</p>
Exemples	APEX

4.6.1.11.10 Atribut: Model de servei

Identificador	ENT011-009
Nom formal	Model de servei
Nom curt	Model de servei
Domini	<p>Aquest atribut recull el model de servei a través del qual l'organització consumeix l'aplicació o sistema. L'atribut combina dos elements conceptuals:</p> <ul style="list-style-type: none"> • El tipus de desenvolupament de l'aplicació o sistema • La propietat de la infraestructura on s'emmagatzema l'aplicació o sistema
Tipus de dades	Alfanumèric categòric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	<p>Aquest atribut ha de prendre un d'aquests valors:</p> <ul style="list-style-type: none"> • Adapt. sol. mercat en servidor altra Adm. • Adapt. sol. mercat en servidor CTTI • Adapt. sol. mercat en servidor extern • Adapt. sol. mercat en servidor propi CEB • Desenv. a mida en servidor altra Adm. • Desenv. a mida en servidor CTTI • Desenv. a mida en servidor extern • Desenv. a mida en servidor propi CEB • Software as a Service (corporatiu) • Software as a Service (privat) • Software as a Service (propi CEB) • Solució de mercat en servidor altra Adm. • Solució de mercat en servidor CTTI • Solució de mercat en servidor extern

	<ul style="list-style-type: none"> • Solució de mercat en servidor propi CEB
Valor per defecte	No aplica.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	Software as a Service (privat)

4.6.1.11.11 Atribut: Abast

Identificador	ENT011-010
Nom formal	Abast
Nom curt	Abast
Domini	<p>Aquest atribut defineix l'àmbit en el qual opera el sistema, i en el qual es prenen decisions sobre la forma i funcionament del sistema. En aquest sentit, es defineix l'abast com:</p> <ul style="list-style-type: none"> • Abast d'usuari final, quan el sistema està desenvolupat pel mateix equip de persones que en serà usuari, i evoluciona únicament d'acord amb les necessitats o conveniència d'aquest equip. • Abast departamental, quan el sistema es desenvolupa per donar suport al funcionament d'un departament o unitat orgànica dins l'organització, però no ha estat desenvolupat per aquest àmbit sinó per un servei professional de desenvolupament tecnològic. • Abast corporatiu, quan el sistema està desenvolupat per un servei professional de desenvolupament tecnològic, sovint l'àrea de tecnologies de la informació i les comunicacions de l'organització, i cobreix les necessitats funcionals de diversos departaments o unitats orgàniques de l'organització, servint per automatitzar processos amb caràcter transversal i amb impacte sobre el conjunt de l'organització.
Tipus de dades	Alfanumèric categòric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	<p>L'atribut pot prendre tres valors:</p> <ul style="list-style-type: none"> • Usuari final • Departamental • Corporatiu <p>Els valors Departamental i Corporatiu es poden descompondre, si cal, en els diferents ens o organismes a qui es vol atribuir responsabilitat sobre el sistema.</p>
Valor per defecte	No aplica.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	Corporatiu

4.6.1.11.12 Atribut: Restriccions semàntiques

Identificador	ENT011-011
Nom formal	Restriccions semàntiques
Nom curt	Restriccions semàntiques
Domini	Descripció de les possibles restriccions o regles a nivell semàntic que puguin afectar a la forma d'emmagatzemar la informació en el sistema.
Tipus de dades	Alfanumèric descriptiu
Mín. ocurrencies	0
Màx. ocurrencies	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	DCAT disposa de l'element de <i>Dublin Core</i> dct:conformsTo per indicar que un recurs ha de ser conforme a un determinat estàndard.
Exemples	Les metadades descriptives de l'objecte " <i>folder</i> " hauran d'estar normalitzades d'acord amb la norma tècnica d'expedient de l'Esquema Nacional d'Interoperabilitat.

4.6.1.11.13 Atribut: Restriccions de format

Identificador	ENT011-012
Nom formal	Restriccions de format
Nom curt	Restriccions de format
Domini	Descripció de les possibles restriccions o regles a nivell de format que puguin afectar a la forma d'emmagatzemar la informació en el sistema.
Tipus de dades	Alfanumèric descriptiu
Mín. ocurrencies	0
Màx. ocurrencies	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	DCAT proposa l'element de <i>Dublin Core</i> dct:format per descriure el format i característiques físiques del recurs o del document.
Exemples	Les dades s'hauran d'emmagatzemar en format delimitat per comes.

4.6.1.11.14 Atribut: Controls d'integritat

Identificador	ENT011-013
Nom formal	Controls d'integritat
Nom curt	Controls d'integritat
Domini	Indicació de l'existència en el sistema de controls que permetin assegurar i/o verificar la integritat de la informació que conté (documents i dades). Aquests controls poden consistir en sistemes de <i>logging</i> o registres d'auditoria, sistemes de signatura o segellat electrònic, sistemes basats en <i>Blockchain</i> , o qualsevol altre sistema

	que asseguri que la informació no es modifica, o en quines condicions ha estat modificada, i que fixi d'aquesta manera la cadena de custòdia.
Tipus de dades	Booleà
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	L'atribut pot prendre dos valors: <ul style="list-style-type: none"> • Sí • No
Valor per defecte	"No"
Fonts	De forma més imprecisa, atès que RiC ho planteja com una informació més pròpia d'instàncies concretes, es pot vincular a l'atribut RiC-A24 Integrity .
Exemples	No

4.6.1.11.15 Atribut: Controls d'accés

Identificador	ENT011-014
Nom formal	Controls d'accés
Nom curt	Controls d'accés
Domini	Indicació de l'existència en el sistema de controls o sistemes d'autenticació que permetin identificar i verificar les identitats que accedeixen al sistema, i assignar-los conjunts de privilegis o capacitats d'actuació dins el sistema.
Tipus de dades	Booleà
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	L'atribut pot prendre dos valors: <ul style="list-style-type: none"> • Sí • No
Valor per defecte	"No"
Fonts	De forma més detallada, i malgrat a RiC és aplicable només a l'entitat RiC-E02 Record resource i derivats, es pot considerar equivalent a l'atribut RiC-A03 Authenticity note , que descriu de quina forma s'autentifica una unitat de descripció, i en conseqüència els usuaris que hi accedeixen. Malgrat en aquest atribut no es contempla el detall de la gestió de permisos, aquest sí que es podria recollir a RiC amb els atributs RiC-A08 Conditions of access , i RiC-A09 Conditions of use . DCAT contempla un element de <i>Dublin Core</i> , dct:accessRights , que correspon a una indicació flexible que pot indicar l'existència d'uns sistema de seguretat en el control de l'accés, com també els agents que poden accedir a un recurs.
Exemples	No

4.6.1.11.16 Atribut: Classificació de seguretat

Identificador	ENT011-015
Nom formal	Classificació segons l'Esquema Nacional de Seguretat
Nom curt	Classificació de seguretat
Domini	<p>Correspon a la classificació per nivells de seguretat o sensibilitat de la informació que estableix per a totes les Administracions públiques espanyoles l'Esquema Nacional de Seguretat, regulat en la seva darrera versió pel Reial Decret 311/2022, de 3 de maig.</p> <p>Aquest atribut s'ha inclòs en el model com a resultat de necessitats específiques de gestió de la informació del Consorci d'Educació de Barcelona. No es considera determinant als efectes d'anàlisi de les dades recollides.</p>
Tipus de dades	Alfanumèric categòric
Mín. ocurrences	1
Màx. ocurrences	1
Regles de validació	<p>L'atribut pot prendre un dels següents valors:</p> <ul style="list-style-type: none">• Públic: sistema que pot ser accedit de forma externa a l'entitat.• Intern: sistema que només pot ser accessible per a usuaris de la pròpia entitat.• Sensible: sistema que només pot ser accessible per a determinats usuaris de la pròpia entitat amb certes restriccions pel que fa a l'accés i la traçabilitat d'aquests accessos.• Crític: sistema que només pot ser accessible per a determinats usuaris de la pròpia entitat amb fortes limitacions en l'ús i procediments especials de seguretat i garantia de disponibilitat.
Valor per defecte	No aplica.
Fonts	<p>Probablement es pot posar en relació amb l'element dct:accessRights, i també amb odrl:hasPolicy del model DCAT.</p> <p>Aquest atribut és equivalent a la metadada Security Classification del model ONDC, i també a la metadada Sensitive Data.</p>
Exemples	Sensible

4.6.1.11.17 Atribut: Descripció

Identificador	ENT011-016
Nom formal	Descripció
Nom curt	Descripció
Domini	Descripció del context de creació i utilització del sistema, de la seva finalitat i del perfil dels seus usuaris. També pot contenir altres informacions relatives a l'ús del sistema.

Tipus de dades	Alfanumèric descriptiu
Mín. ocurrencies	0
Màx. ocurrencies	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	<p>Aquest atribut equival a la propietat 5 Description del model AGRkMS.</p> <p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE5 - Descripción.</p> <p>RiC recull el concepte de descripció o comentaris de caràcter genèric en l'atribut RiC-A16 Descriptive note.</p> <p>En el cas de bases de dades relacionals és equiparable a la metadada M_5.1-1 description de SIARD.</p> <p>DCAT inclou l'element de Dublin Core dct:description per a descriure el recurs que conté les dades.</p> <p>Aquest atribut és equivalent a la metadada Description del model ONDC.</p>
Exemples	Aquesta aplicació s'utilitza per al registre de l'alumnat amb necessitats específiques de suport educatiu pel que fa a la seva situació econòmica i social. Es va desenvolupar com a resultat de la resolució de reconeixement de NESE-B, i és utilitzat preferentment per les treballadores socials de l'entitat.

4.6.1.11.18 Atribut: Funcionalitats

Identificador	ENT011-017
Nom formal	Funcionalitats
Nom curt	Funcionalitats
Domini	Descripció de les funcionalitats bàsiques del sistema, de la seva estructura interna, de la seva interacció amb d'altres sistemes, i d'altres informacions de caràcter tecnològic o d'implementació que es considerin pertinents.
Tipus de dades	Alfanumèric descriptiu
Mín. ocurrencies	0
Màx. ocurrencies	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	De forma indirecta pot tenir relació amb la informació consignada en l'atribut RiC-A41 Technical characteristics .
Exemples	Aquesta aplicació permet l'alta, modificació i baixa d'alumnes, i està connectada amb el registre d'alumnat de Catalunya, contra el qual valida la informació de matrícula de l'alumnat.

4.6.1.11.19 Atribut: Ubicació

Identificador	ENT011-018
Nom formal	Ubicació
Nom curt	Ubicació
Domini	<p>Ubicació física del sistema dins la xarxa de comunicacions de l'entitat, als efectes de poder establir connexions i interoperar amb el sistema. Pot incloure elements com URL de components del sistema, adreces de xarxa local, referències de vistes de base de dades, etc.</p> <p>Aquest atribut s'ha inclòs en el model com a resultat de necessitats específiques de gestió de la informació del Consorci d'Educació de Barcelona. No es considera determinant als efectes d'anàlisi de les dades recollides.</p>
Tipus de dades	Alfanumèric
Mín. ocurrencies	0
Màx. ocurrencies	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	<p>DCAT disposa de l'element dcat:accessURL per indicar la URL que dona accés a un determinat recurs (<i>dataset</i> o <i>distribution</i>). També disposa d'un altre element en el cas de recursos que es puguin descarregar, l'element dcat:downloadURL. I pel que fa als <i>data services</i>, disposa també de l'element dcat:endpointURL per indicar l'arrel de la ubicació del servei.</p> <p>Aquest atribut és equivalent a la metadada Access URL del model ONDC.</p>
Exemples	https://gencat.sharepoint.com/sites/serveiorientacio

4.6.1.11.20 Atribut: Data inici servei

Identificador	ENT011-019
Nom formal	Data d'inici del servei
Nom curt	Data inici servei
Domini	<p>La data d'inici o entrada en servei d'un sistema d'informació és la data en què la primera versió del sistema puja a producció i es posa en funcionament entre els que seran els seus usuaris habituals. La data de pujada a producció és una data concreta que és possible fixar amb prou precisió.</p>
Tipus de dades	Data
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	Data vàlida.
Valor per defecte	No aplica.
Fonts	<p>Aquest atribut equival a la propietat 4 Date range del model AGRkMS.</p>

	<p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE4.1 - Fecha inicio.</p> <p>Equival a l'atribut RiC-A29 Normalized date dins l'entitat RiC-E20 Date range. Vegeu nota explicativa a l'apartat Fonts de l'atribut ENT001-006.</p> <p>DCAT proposa l'ús de l'element de Dublin Core dct:issued per indicar la data en què el recurs ha començat a estar disponible.</p> <p>Aquest atribut és equivalent a la metadada Temporal coverage from del model ONDC.</p>
Exemples	12 de gener de 2008.

4.6.1.11.21 Atribut: Data fi servei

Identificador	ENT011-020
Nom formal	Data de fi de servei
Nom curt	Data fi servei
Domini	La data de fi de servei és la data en què un sistema d'informació deixa de funcionar de forma definitiva, i deixa d'estar disponible en un entorn productiu. Els canvis de versió no determinen aquesta data, només la determina la finalització definitiva de la disponibilitat del sistema. Aquesta indisponibilitat només afecta les funcionalitats, no la informació continguda en el sistema, que pot continuar operativa un cop migrada a un nou sistema, o bé estabilitzada dins el sistema finalitzat en forma de dades llegades (<i>legacy system</i>).
Tipus de dades	Data
Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	Data vàlida.
Valor per defecte	No aplica.
Fonts	<p>Aquest atribut equival a la propietat 4 Date range del model AGRkMS.</p> <p>El model e-EMGDE contempla per aquest atribut la metadada eEMGDE4.2 - Fecha fin.</p> <p>Equival a l'atribut RiC-A29 Normalized date dins l'entitat RiC-E20 Date range. Vegeu nota explicativa a l'apartat Fonts de l'atribut ENT001-006.</p> <p>Es podria entendre equiparable, en el cas de bases de dades relacionals, amb la metadada M_5.1-1 archivalDate de SIARD.</p> <p>Aquest atribut és equivalent a la metadada Temporal coverage to del model ONDC.</p>
Exemples	31 de desembre de 2021.

4.6.1.12 Entitat: Interfície de connexió entre sistemes

4.6.1.12.1 Descripció de l'entitat

Identificador	ENT012
Nom formal	Interfície de connexió entre sistemes
Nom curt	Interfície
Domini	<p>En el context d'aquest model anomenem interfície a qualsevol flux de dades entre dos sistemes d'informació. La interfície pot incorporar algun element tecnològic (per exemple un connector, el codi d'un procés ETL, etc.), o bé limitar-se a un procés d'exportació-importació de dades. En tota interfície sempre s'indica un sistema que és origen de les dades, un sistema destí, unes dades que es transfereixen i unes condicions d'aquesta transferència.</p> <p>El concepte d'interfície d'aquest model correspon en l'àmbit de la ciència de dades al concepte de flux de dades (<i>dataflow</i>).</p>
Atributs	<p>ENT012-001 Identificador ENT012-002 Nom ENT012-003 Tipus d'interfície ENT012-004 Tipus d'importació ENT012-005 Procediment d'importació ENT012-006 Actualització ENT012-007 Historificació ENT012-008 Nom del fitxer de càrrega ENT012-009 Restriccions semàntiques ENT012-010 Restriccions de format ENT012-011 Observacions ENT012-012 Data inici servei ENT012-013 Data fi servei</p>
Relacions	<p>REL022 Pot traslladar una o més dades elementals</p> <p>ENT012 Pot comunicar un sistema d'informació (origen) amb un sistema d'informació (destí)</p>
Fonts	<p>DCAT ofereix dues possibles connexions amb el concepte d'interfície d'aquest model. La primera és a través del concepte de <i>Data service</i>, un dels dos elements que en el model DCAT es poden incloure dins les descripcions que conté un catàleg de dades. El <i>Data service</i> correspon a qualsevol tipus d'operació, accessible mitjançant una interfícies (API), que proveeix accés a un o més conjunts de dades o funcions de processament de dades. A DCAT es representa amb la classe dcat:DataService.</p> <p>DCAT ofereix una altra opció equivalent al concepte d'interfície d'aquest model. Es tracta de l'element dcat:qualifiedRelation, que correspon a qualsevol mena de relació entre recursos. També ho contempla a través de la classe dcat:Relationship.</p> <p>El model ONDC, tot i que no descriu relacions, sí que inclou algunes elements descriptius propis d'aquesta entitat.</p>

4.6.1.12.2 Atribut: Identificador (codi)

Identificador	ENT012-001
Nom formal	Identificador
Nom curt	ID
Domini	Codi que identifica de forma unívoca la interfície dins de l'organització.
Tipus de dades	Alfanumèric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	El codi de la interfície és un codi qualificat compost dels següents elements: <ul style="list-style-type: none">• Inicials "IF"• Número seqüencial de quatre dígit
Valor per defecte	No aplica.
Fonts	Correspon a l'atribut genèric RiC-A22 Identifier . DCAT utilitza l'element de <i>Dublin Core</i> dct:identifier per assignar un identificador únic a cada recurs, i també a cada servei de dades.
Exemples	DF0017

4.6.1.12.3 Atribut: Nom

Identificador	ENT012-002
Nom formal	Nom
Nom curt	Nom
Domini	Nom comú que se li dona a la interfície, als efectes d'identificació ràpida en l'ús habitual d'aquesta descripció.
Tipus de dades	Alfanumèric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	Correspon a l'atribut genèric RiC-A28 Name . DCAT proposa l'element de <i>Dublin Core</i> dct:title per indicar el nom del recurs que conté les dades, i també del servei de dades.
Exemples	Càrrega diària dels llistats de GEDAC

4.6.1.12.4 Atribut: Tipus d'interfície

Identificador	ENT012-003
Nom formal	Tipus d'interfície
Nom curt	Tipus d'interfície
Domini	Tipus d'interfície categoritzada en funció del nivell d'intervenció humana en l'execució del flux de dades.
Tipus de dades	Alfanumèric categòric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	L'atribut pot prendre tres valors: <ul style="list-style-type: none">• Automatitzada: tot el flux de dades s'executa sense cap intervenció humana, de forma desatesa.• Accés a aplicació: el flux de dades requereix l'accés a l'aplicació origen per part d'un usuari, que pot ser humà o equivalent (un procés RPA emulant accions humanes).• Obtenció manual de fitxer: el flux de dades s'executa a base d'una seqüència d'accions realitzades per un agent humà.
Valor per defecte	No aplica.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	Automatitzada

4.6.1.12.5 Atribut: Tipus d'importació

Identificador	ENT012-004
Nom formal	Tipus d'importació
Nom curt	Tipus d'importació
Domini	Forma com es porta a terme l'execució del flux de dades, en funció de la quantitat d'operacions realitzades i del sistema més actiu (origen o destí).
Tipus de dades	Alfanumèric categòric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	L'atribut pot prendre tres valors: <ul style="list-style-type: none">• Hostatjat en aplicació origen: les dades resideixen en el sistema origen, i són accedides pel sistema destí quan n'ha de menester (sistema <i>pull</i>).• Flux de dades entre sistemes: el sistema origen genera en un determinat moment una extracció de dades que envia al sistema destí, qui procedeix a la seva càrrega (sistema <i>push</i>).• Càrrega de fitxer independent: un agent extern, habitualment humà, descarrega les dades del sistema origen i les carrega al sistema destí.
Valor per defecte	No aplica.

Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	Flux de dades entre sistemes

4.6.1.12.6 Atribut: Procediment d'importació

Identificador	ENT012-005
Nom formal	Procediment d'importació
Nom curt	Procediment d'importació
Domini	El procediment d'importació fixa el nivell d'intervenció humana o d'un agent extern en el flux de dades entre els sistemes origen i destí.
Tipus de dades	Alfanumèric categòric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	<p>L'atribut pot prendre quatre valors, en funció del nivell d'intervenció humana:</p> <ul style="list-style-type: none"> • Manual amb depuració prèvia: un agent humà intervé en totes les etapes del flux de dades, incloent la descàrrega del sistema origen, el tractament i preparació de les dades, i la càrrega al sistema destí. • Manual directe: l'agent humà intervé en les accions de descàrrega i càrrega, actuant com a disparador, però la seva acció, en no existir necessitat de tractament a mida de les dades, es pot combinar amb processos automàtics de processament de les dades en el moment de la càrrega. • Robotitzat: l'acció humana del procés anterior és substituïda per un robot administratiu (RPA). • Connexió directa a origen: els sistemes origen i destí estan directament connectats, i disposen cadascun dels procediments necessaris per a disparar la descàrrega, preparació de les dades i càrrega a destí.
Valor per defecte	No aplica.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	Robotitzat

4.6.1.12.7 Atribut: Actualització

Identificador	ENT012-006
Nom formal	Període d'actualització
Nom curt	Actualització
Domini	Període que transcorre de forma regular entre una execució del flux de dades i la següent.
Tipus de dades	Alfanumèric categòric
Mín. ocurrencies	1

Màx. ocurrencies	1
Regles de validació	L'atribut pot prendre cinc valors: <ul style="list-style-type: none"> • Anual • Mensual • Setmanal • Diària • A demanda
Valor per defecte	No aplica.
Fonts	DCAT contempla l'element de <i>Dublin Core</i> dct:accrualPeriodicity per indicar la freqüència amb la qual un <i>dataset</i> és actualitzat i/o publicat. Aquest atribut és equivalent a la metadada Update frequency del model ONDC .
Exemples	Setmanal

4.6.1.12.8 Atribut: Historificació

Identificador	ENT012-007
Nom formal	Període d'historificació
Nom curt	Historificació
Domini	Període que transcorre entre la generació d'una versió d'arxiu de les dades i la següent. En el cas que el sistema destí no separi les dades entre dades vigents i versions d'arxiu, direm que la historificació és acumulativa.
Tipus de dades	Alfanumèric categòric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	L'atribut pot prendre tres valors: <ul style="list-style-type: none"> • Anual • Mensual • Acumulativa
Valor per defecte	No aplica.
Fonts	DCAT inclou l'element dcat:temporalResolution per indicar el temps entre captures d'un <i>dataset</i> en sèries temporals. També proposa l'ús de l'element de <i>Dublin Core</i> dct:temporal per indicar el termini de temps cobert per un <i>dataset</i> .
Exemples	Anual

4.6.1.12.9 Atribut: Nom del fitxer de càrrega

Identificador	ENT012-008
Nom formal	Nom del fitxer de càrrega
Nom curt	Nom del fitxer de càrrega
Domini	<p>Nom assignat al fitxer que genera la descàrrega del sistema origen, i que posteriorment es carrega al sistema destí. S'informarà preferentment en les importacions manuals de fitxers independents. La descàrrega pot generar més d'un fitxer.</p> <p>Aquest atribut s'ha inclòs en el model com a resultat de necessitats específiques de gestió de la informació del Consorci d'Educació de Barcelona. No es considera determinant als efectes d'anàlisi de les dades recollides.</p>
Tipus de dades	Alfanumèric
Mín. ocurrències	0
Màx. ocurrències	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	DCAT disposa d'un element en el cas de recursos que es puguin descarregar, l'element dcat:downloadURL . No és el cas exacte d'aquest atribut, llevat que s'hi indiqui el <i>path</i> complet del fitxer.
Exemples	Dades_curs_21-22.csv

4.6.1.12.10 Atribut: Restriccions semàntiques (dades)

Identificador	ENT012-009
Nom formal	Restriccions semàntiques
Nom curt	Restriccions semàntiques
Domini	Qualsevol tipus de condicionant pel que fa a la representació i significació de les dades que es descarreguen del sistema origen, i que pugui determinar aspectes de la seva importació al sistema destí.
Tipus de dades	Alfanumèric descriptiu
Mín. ocurrències	0
Màx. ocurrències	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	DCAT disposa de l'element de <i>Dublin Core</i> dct:conformsTo per indicar que un recurs ha de ser conforme a un determinat estàndard.
Exemples	Les dades relatives al país (columna 12) han d'estar codificades d'acord amb els codis de país de la norma ISO 3166-1 alpha3.

4.6.1.12.11 Atribut: Restriccions de format

Identificador	ENT012-010
Nom formal	Restriccions de format
Nom curt	Restriccions de format
Domini	Qualsevol tipus de condicionant pel que fa a la codificació de les dades que es descarreguen del sistema origen, i que pugui determinar aspectes de la seva importació al sistema destí.
Tipus de dades	Alfanumèric descriptiu
Mín. ocurrences	0
Màx. ocurrences	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	DCAT proposa l'element de <i>Dublin Core</i> dct:format per descriure el format i característiques físiques del recurs o del servei de dades.
Exemples	Les dades han d'estar en format de valors separats per comes (CSV), d'acord amb el tipus MIME text/csv (PUID x-fmt/18).

4.6.1.12.12 Atribut: Observacions

Identificador	ENT012-011
Nom formal	Observacions
Nom curt	Observacions
Domini	Comentaris sobre qualsevol aspecte del funcionament de la interfície no recollit en cap dels atributs anteriors.
Tipus de dades	Alfanumèric descriptiu
Mín. ocurrences	0
Màx. ocurrences	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	RiC recull el concepte de descripció o comentaris de caràcter genèric en l'atribut RiC-A16 Descriptive note .
Exemples	Es preveu el desenvolupament d'un nou connector que substituirà l'actual dins l'inici de l'any vinent.

4.6.1.12.13 Atribut: Data inici servei

Identificador	ENT012-012
Nom formal	Data d'inici del servei
Nom curt	Data inici servei
Domini	La data d'inici o entrada en servei d'un flux de dades és la data en què aquest flux, un cop definit, s'executa per primera vegada entre sistemes d'informació que estan en producció.
Tipus de dades	Data
Mín. ocurrences	1
Màx. ocurrences	1

Regles de validació	Data vàlida.
Valor per defecte	No aplica.
Fonts	Equival a l'atribut RiC-A29 Normalized date dins l'entitat RiC-E20 Date range . Vegeu nota explicativa a l'apartat Fonts de l'atribut ENT001-006 . DCAT proposa l'ús de l'element de <i>Dublin Core</i> dct:issued per indicar la data en què el servei de dades ha començat a estar disponible.
Exemples	12 de gener de 2008.

4.6.1.12.14 Atribut: Data fi servei

Identificador	ENT012-013
Nom formal	Data de fi de servei
Nom curt	Data fi servei
Domini	La data de fi de servei és la data en què es considera de forma convencional que el flux es deixarà d'executar en la forma en què va estar definit, o bé es desmantellen els seus components. Una modificació d'aspectes semàntics o de format del flux de dades no es considerarà una finalització del servei.
Tipus de dades	Data
Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	Data vàlida.
Valor per defecte	No aplica.
Fonts	Equival a l'atribut RiC-A29 Normalized date dins l'entitat RiC-E20 Date range . Vegeu nota explicativa a l'apartat Fonts de l'atribut ENT001-006 .
Exemples	31 de desembre de 2021.

4.6.1.13 Entitat: Visualització de dades

4.6.1.13.1 Descripció de l'entitat

Identificador	ENT013
Nom formal	Visualització de dades
Nom curt	Visualització de dades
Domini	Instrument o producte d'informació a través del qual es pot accedir a diferents visualitzacions textuales, numèriques o gràfiques de conjunts de dades combinats. Habitualment s'anomena <i>dashboard</i> o quadre de comandament, i s'utilitza tant per a finalitats de suport a la gestió, com de presa de decisions (estratègiques i operatives) i anàlisi estadística de situacions i polítiques.
Atributs	ENT013-001 Identificador ENT013-002 Nom ENT013-003 Tipus ENT013-004 Finalitat ENT013-005 Destinataris ENT013-006 Usuaris ENT013-007 Data inici servei ENT013-008 Data fi servei
Relacions	REL027 Conté una o més dades elementals
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.

4.6.1.13.2 Atribut: Identificador (codi)

Identificador	ENT013-001
Nom formal	Identificador
Nom curt	ID
Domini	Codi que identifica de forma unívoca cada visualització de dades.
Tipus de dades	Alfanumèric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	Es tracta d'un codi qualificat que es compon dels següents elements: <ul style="list-style-type: none">• Lletres QC• Número seqüencial de tres dígitos que identifica el quadre• Guió• Número seqüencial de dos dígitos que identifica la versió del quadre
Valor per defecte	No aplica.
Fonts	Correspon a l'atribut genèric RiC-A22 Identifier .
Exemples	QC066-02

4.6.1.13.3 Atribut: Nom

Identificador	ENT013-002
Nom formal	Nom
Nom curt	Nom
Domini	Nom assignat a la visualització de dades.
Tipus de dades	Alfanumèric
Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	Correspon a l'atribut genèric RiC-A28 Name .
Exemples	Quadre de seguiment de l'alumnat vulnerable

4.6.1.13.4 Atribut: Tipus

Identificador	ENT013-003
Nom formal	Tipus de visualització
Nom curt	Tipus
Domini	Tipus de visualització segons el seu context i moment d'utilització.
Tipus de dades	Alfanumèric categòric
Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	Aquest atribut pot prendre dos valors: <ul style="list-style-type: none">• De gestió: quan la visualització s'utilitza de forma simultània a la gestió dels processos i amb l'objectiu de prendre-hi decisions bàsicament operatives.• Estadística: quan la visualització s'utilitza de forma posterior a l'execució dels processos i amb l'objectiu de prendre decisions de caràcter estratègic (avaluació i/o planificació).
Valor per defecte	No aplica.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	De gestió

4.6.1.13.5 Atribut: Finalitat

Identificador	ENT013-004
Nom formal	Finalitat
Nom curt	Finalitat
Domini	Descripció textual de la finalitat per a la qual s'ha construït la visualització, i del detall del seu context d'utilització. Pot fer

	<p>referència als processos als quals dona suport, i també al tipus de decisions que pot ajudar a prendre.</p> <p>Aquest atribut s'ha inclòs en el model com a resultat de necessitats específiques de gestió de la informació del Consorci d'Educació de Barcelona. No es considera determinant als efectes d'anàlisi de les dades recollides.</p>
Tipus de dades	Alfanumèric descriptiu
Mín. ocurrències	0
Màx. ocurrències	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	Aquesta visualització té per objectiu facilitar els indicadors sobre equitat en la distribució de l'alumnat pels diferents centres educatius necessaris per a la planificació del pla de xoc contra la segregació escolar del curs vinent.

4.6.1.13.6 Atribut: Destinataris

Identificador	ENT013-005
Nom formal	Destinataris
Nom curt	Destinataris
Domini	<p>Descripció dels diferents perfils d'usuari que accedeixen a la visualització de dades.</p> <p>Aquest atribut s'ha inclòs en el model com a resultat de necessitats específiques de gestió de la informació del Consorci d'Educació de Barcelona. No es considera determinant als efectes d'anàlisi de les dades recollides.</p>
Tipus de dades	Alfanumèric descriptiu
Mín. ocurrències	0
Màx. ocurrències	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	Tècnics de gestió del Pla de Xoc contra la segregació escolar, i treballadors/ores socials dels serveis educatius.

4.6.1.13.7 Atribut: Usuaris

Identificador	ENT013-006
Nom formal	Usuaris
Nom curt	Usuaris
Domini	Relació nominal de persones que accedeixen a la visualització de dades. No es considera un tipus de relació entre agent i sistema, atès

	<p>que els agents sempre corresponen a identitats col·lectives, o bé a rols, i no a individus concrets.</p> <p>Aquest atribut s'ha inclòs en el model com a resultat de necessitats específiques de gestió de la informació del Consorci d'Educació de Barcelona. No es considera determinant als efectes d'anàlisi de les dades recollides.</p>
Tipus de dades	Usuari
Mín. ocurrencies	0
Màx. ocurrencies	100
Regles de validació	Aquest atribut es valida contra el directori corporatiu d'usuaris de l'organisme amb llicència d'accés a visualitzacions de dades.
Valor per defecte	No aplica.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	08451478@xtec.cat

4.6.1.13.8 Atribut: *Data inici servei*

Identificador	ENT013-007
Nom formal	Data d'inici del servei
Nom curt	Data inici servei
Domini	La data de posada en producció de la visualització de dades, quan es publica per primer cop i es fa accessible a usuaris.
Tipus de dades	Data
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	Data vàlida.
Valor per defecte	No aplica.
Fonts	Equival a l'atribut RiC-A29 Normalized date dins l'entitat RiC-E20 Date range . Vegeu nota explicativa a l'apartat Fonts de l'atribut ENT001-006 .
Exemples	12 de gener de 2008.

4.6.1.13.9 Atribut: *Data fi servei*

Identificador	ENT013-008
Nom formal	Data de fi de servei
Nom curt	Data fi servei
Domini	La data en què la visualització es despublica i deixa d'estar accessible als seus usuaris en la seva forma definitiva. Això no implica l'eliminació de la visualització, atès que un cop donat de baixa s'historifica i pot servir d'antecedent o suport a futurs visualitzacions, tractaments de dades o anàlisi estadístiques en qualitat de versió d'arxiu.
Tipus de dades	Data
Mín. ocurrencies	1

Màx. ocurrències	1
Regles de validació	Data vàlida.
Valor per defecte	No aplica.
Fonts	Equival a l'atribut RiC-A29 Normalized date dins l'entitat RiC-E20 Date range . Vegeu nota explicativa a l'apartat Fonts de l'atribut ENT001-006 .
Exemples	31 de desembre de 2021.

4.6.1.14 Entitat: Objecte

4.6.1.14.1 Descripció de l'entitat

Identificador	ENT014
Nom formal	Objecte
Nom curt	Objecte
Domini	Cosa o acció del món real ²²¹ que pot ser descrita o caracteritzada per una o més dades.
Atributs	ENT014-001 Nom ENT014-002 Tipus
Relacions	REL024 Pot ser unitat de descripció d'un o més sistemes d'informació REL025 Pot ser subtipus d'un objecte REL025 Pot ser supertipus d'un o més objectes REL026 És descrit per una o més dades elementals
Fonts	<p>El model NEDA és un dels pocs models arxivístics que, anticipant a RiC, contempla una entitat per a descriure coses del món real, l'entitat que anomena Concepto, objeto o acontecimiento.</p> <p>RiC contempla una entitat que anomena RiC-E01 Thing, la qual descriu qualsevol cosa del domini de l'experiència humana que pot ser descrita. D'acord amb aquesta definició, podríem establir una correspondència entre l'atribut Thing de RiC i aquesta entitat. Però hem de parar esment que RiC utilitza l'entitat Thing per a definir el nivell més alt de descripció dins l'ontologia RiC, amb la qual cosa hem d'interpretar que inclou la resta d'entitats del model. Des d'aquest punt de vista, RiC permetria caracteritzar els tipus d'objecte descrit d'acord amb les entitats de segon nivell de RiC, de forma similar a com ho fa l'atribut ENT014-002 Tipus d'objecte.</p> <p>DCAT inclou diverses opcions per classificar <i>datasets</i> i assignar-los diferents categories. Ho fa través de l'element dcat:themeTaxonomy, i ho combina amb el vocabulari SKOS a través de les classes skos:ConceptScheme i skos:Concept.</p> <p>Dues de les tres entitats principals del model PROV poden ser considerades equivalents a aquesta entitat. La primera és l'entitat prov:Activity, que en aquest cas ofereix una doble equivalència: amb els objectes de tipus "Acció" (informat a l'atribut ENT014-002 Tipus d'objecte), i amb el concepte de procés (entitat ENT002 Procés). I la segona és l'entitat prov:Entity, que el model PROV defineix com "<i>a physical, digital, conceptual, or other kind of thing with some fixed aspects; entities may be real or imaginary</i>". Aquesta segona entitat es pot considerar equivalent als objectes de tipus "Cosa" (informat a l'atribut ENT014-002 Tipus d'objecte). L'entitat prov:Entity es desplega en tres subclasses: prov:Collection, prov:Bundle, i</p>

²²¹ En terminologia **RiC**, cosa o acció "del domini de l'experiència humana".

	prov:Plan. De tots els models de descripció analitzats, PROV és el que més clarament encaixa en l'essència de l'entitat objecte d'aquest model.
--	---

4.6.1.14.2 Atribut: Nom

Identificador	ENT014-001
Nom formal	Nom
Nom curt	Nom
Domini	Nom assignat a l'objecte. Habitualment s'assigna el nom en el moment en què s'identifiquen les dades, i es relata en llenguatge natural allò que és descrit per la dada.
Tipus de dades	Alfanumèric
Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	No
Valor per defecte	No aplica.
Fonts	Correspon a l'atribut genèric RiC-A28 Name .
Exemples	Centre educatiu

4.6.1.14.3 Atribut: Tipus

Identificador	ENT014-002
Nom formal	Tipus d'objecte
Nom curt	Tipus
Domini	Tipus d'objecte segons el seu caràcter actiu o passiu. Als efectes d'anàlisi, es caracteritzarà l'objecte com a cosa, quan es defineixi preferentment per un substantiu, i com a acció quan es defineixi amb un verb substantivat.
Tipus de dades	Alfanumèric categòric
Mín. ocurrències	1
Màx. ocurrències	1
Regles de validació	Aquest atribut pot prendre dos valors: <ul style="list-style-type: none"> • Cosa: qualsevol cosa del món real que pot ser descrita. • Acció: qualsevol acció, esdeveniment, tràmit o procés que té lloc en el món real i que pot ser descrit.
Valor per defecte	No aplica.
Fonts	Els valors d'aquest atribut són equivalents a la jerarquia d'entitats del model RiC (utilitzant per exemple RiC-E01 Thing per a cosa, RiC-E15 Activity per acció, etc.).
Exemples	Cosa

4.6.2 Relacions

4.6.2.1 Relació: agent amb agent

4.6.2.1.1 Descripció de la relació

Identificador	REL001
Domini	ENT001 Agent
Rang	ENT001 Agent
Significat de la relació	Depèn/És staff de
Relació inversa	Depèn/És staff de
Obligatorietat	No
Cardinalitat	N a N
Definició	Defineix el tipus de relació que mantenen dos agents entre ells. Habitualment es correspon amb relacions orgàniques o jeràrquiques.
Atributs	REL001-001 Rol de l'agent
Fonts	<p>Aquesta relació es pot considerar equivalent a l'element 5.2.7 Internal structures/Genealogy de la norma ISAAR. I també queda reflectida amb 5.3.1 Names/Identifiers of related corporate bodies, persons or families.</p> <p>Correspon a la relació RiC-R045 Agent is subordinate of i inversa, i també pot correspondre a la relació RiC-R041 Agent is controller of Agent i inversa. A RiC es desenvolupa també com a relació de conseqüència amb RiC-R016 Agent is successor of Agent i inversa.</p> <p>És equivalent a la relació Agente - Agente del model NEDA.</p>

4.6.2.1.2 Atribut: Rol de l'agent

Identificador	REL001-001
Nom formal	Rol de l'agent
Nom curt	Rol de l'agent
Domini	Defineix el tipus de relació que mantenen dos agents entre ells, habitualment segons un patró de dependència jeràrquica.
Tipus de dades	Alfanumèric categòric
Mín. ocurrences	1
Màx. ocurrences	1
Regles de validació	Aquest atribut pot prendre dos valors: <ul style="list-style-type: none">• Depèn de• És staff de
Valor per defecte	No aplica.

Fonts	Aquest atribut es pot considerar equivalent a l'element 5.3.2 Category of relationship de la norma ISAAR .
Exemples	Depèn de

4.6.2.2 Relació: agent amb procés

4.6.2.2.1 Descripció de la relació

Identificador	REL002
Domini	ENT001 Agent
Rang	ENT002 Procés
Significat de la relació	És responsable/destinatari de
Relació inversa	Té com a responsable/destinatari a
Obligatorietat	No/Sí
Cardinalitat	N a N
Definició	Els agents tenen diferents formes d'intervenció en els processos, i poden tenir un paper actiu o passiu. Aquesta relació recull la definició d'aquest rol. La relació té diferent nivell d'obligatorietat: tot procés té un responsable i pot tenir un o més destinataris, mentre que tot agent pot ser responsable o destinatari d'un o més processos.
Atributs	REL002-001 Rol de l'agent
Fonts	És equivalent a la relació Agente - Función y sus divisiones del model NEDA .

4.6.2.2.2 Atribut: Rol de l'agent

Identificador	REL002-001
Nom formal	Rol de l'agent
Nom curt	Rol de l'agent
Domini	<p>Defineix el paper que diferents agents, interns o externs, tenen dins l'execució d'un procés, sigui com agent principal o responsable, o com a destinatari o receptor del resultat del procés.</p> <ul style="list-style-type: none"> Definirem el propietari del procés com aquell agent que determina el desenvolupament del procés o n'assumeix un rol preminent per la competència o pel nombre de tasques assignades. En cas que un procés generi una sola derivació de sèries documentals, aquest agent serà també el productor de la sèrie primària. Definirem destinatari del procés com aquells agents als quals s'adreça el resultat del procés, sigui aquest un producte, un servei o un control.

Tipus de dades	Alfanumèric categòric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	Aquest atribut pot prendre dos valors: <ul style="list-style-type: none"> • Responsable • Destinatari
Valor per defecte	No aplica.
Fonts	MoReq recull aquest atribut en forma d'entitat a través de l'entitat E14.2.13 Role . En el rol de destinatari, RiC contempla l'atribut RiC-A15 Demographic Group per a descriure qualsevol grup humà amb característiques comunes, incloent també la possibilitat de definir perfils de persona o categories similars al rol.
Exemples	Responsable

4.6.2.3 Relació: agent amb tasca

4.6.2.3.1 Descripció de la relació

Identificador	REL003
Domini	ENT001 Agent
Rang	ENT003 Tasca
Significat de la relació	Executor de
Relació inversa	Executada per
Obligatorietat	No/Sí
Cardinalitat	1 a N
Definició	Una tasca és sempre executada per un agent. Aquest agent pot ser individual o col·lectiu, i també automatitzat. En aquest darrer cas, es considera com agent aquell en nom del qual està actuant l'agent automatitzat.
Atributs	Sense atributs propis.
Fonts	MoReq recull aquesta relació amb la metadada M14.4.83 Performed By User Identifier , i també amb M14.4.81 Participating User Identifier , de l'entitat E14.2.8 Event . Correspon a la relació RiC-R060 Activity is performed by Agent i inversa.

4.6.2.4 Relació: agent amb sèrie documental

4.6.2.4.1 Descripció de la relació

Identificador	REL004
Domini	ENT001 Agent
Rang	ENT007 Sèrie documental
Significat de la relació	Productor de
Relació inversa	Produïda per
Obligatorietat	No/Sí
Cardinalitat	1 a N
Definició	Una sèrie documental és el resultat de l'acumulació d'evidències que generen les accions de recepció i arxiu per part d'un agent, que n'esdevé el productor. El concepte de productor, doncs, correspon al de l'agent que reuneix i conserva un conjunt d'evidències com a resultat de la seva participació en un determinat procés.
Atributs	Sense atributs propis.
Fonts	<p>El concepte de productor de la sèrie documental correspon a l'element ISAD 3.2.1 Name of creator(s).</p> <p>Aquest concepte de productor també apareix recollit a EAD amb un element explícit, <origination>, i un que pot actuar d'identificador únic del productor: <agencycode>.</p> <p>És equivalent a la relació Documentos de archivo (excepto colección, división de colección y componente documental) - Agente del model NEDA, on l'agent adquireix el significat de Productor.</p> <p>Correspon a la relació RiC-R026 Agent is provenance of Records resource i inversa.</p> <p>Aquest atribut és equivalent a la metadada Data Custodian del model ONDC.</p>

4.6.2.5 Relació: agent amb regla d'accés

4.6.2.5.1 Descripció de la relació

Identificador	REL005
Domini	ENT001 Agent
Rang	ENT008 Regla d'accés
Significat de la relació	Obté permisos d'accés de

Relació inversa	Atorga permisos d'accés a
Obligatorietat	No/Sí
Cardinalitat	N a N
Definició	Les regles d'accés defineixen les capacitats d'un determinat agent o grup d'usuaris en relació amb un determinat actiu d'informació, que en el nostre model és sempre una sèrie documental. Aquestes capacitats poden implicar no només la visualització, sinó també si escau la modificació o eliminació.
Atributs	Sense atributs propis.
Fonts	Correspon a la relació RiC-R067 Mandate autorizes Agent i inversa .

4.6.2.6 Relació: agent amb sistema d'informació

4.6.2.6.1 Descripció de la relació

Identificador	REL006
Domini	ENT001 Agent
Rang	ENT011 Sistema d'informació
Significat de la relació	Promotor de
Relació inversa	Promogut per
Obligatorietat	No/Sí
Cardinalitat	1 a N
Definició	El primer pas en el desenvolupament d'un sistema d'informació és la definició de requeriments funcionals, que responen habitualment a les necessitats informacionals i d'execució d'accions d'un determinat agent. Aquest agent esdevé en conseqüència el promotor del sistema.
Atributs	Sense atributs propis.
Fonts	El concepte d'agent promotor és equivalent a SIARD M_5.1-1 data Owner , i també al de dct:creator de <i>Dublin Core</i> , usat a DCAT , tot i que també disposa de l'element de Dublin Core dct:publisher per indicar aquell agent que fa disponible un recurs. DCAT també té inclou l'element prov:qualifiedAttribution de l'ontologia PROV , que permet indicar un vincle amb qualsevol agent que tingui alguna forma de responsabilitat sobre el recurs. Aquest atribut és equivalent a la metadada Data Custodian del model ONDC .

4.6.2.7 Relació: agent amb evidència

4.6.2.7.1 Descripció de la relació

Identificador	REL007
Domini	ENT001 Agent
Rang	ENT004 Evidència
Significat de la relació	Crea/Rep
Relació inversa	Creada per/Rebuda per
Obligatorietat	No/Sí
Cardinalitat	2 a N
Definició	<p>Tota evidència és el resultat d'una acció realitzada per un agent, que esdevé el creador de l'evidència, tant si aquesta creació ha estat voluntària com si ha estat un acte de constància automatitzat. Una part de les evidències responen a comunicacions, amb la qual cosa també hi haurà en aquests casos un agent receptor, diferent del creador.</p> <p>Si un document té diversos destinataris, es crearan tantes evidències com destinataris, més una evidència pel remitent. En els casos de creació d'evidències amb la mera finalitat de generar constància d'un fet o acció, s'informarà com agent receptor el mateix agent creador.</p>
Atributs	REL007-001 Rol de l'agent
Fonts	<p>El model e-EMGDE contempla per aquesta relació la metadada eEMGDE24 - Órgano. I també la metadada eEMGDE25 - Origen del documento.</p> <p>És equivalent a la relació Documentos de archivo - Agente del model NEDA.</p> <p>Correspon a la relació RiC-R031 Agent is sender of Records resource i inversa, i també a la relació RiC-R029 Agent is receiver of Records resource i inversa. També es pot considerar equivalent a la relació RiC-R027 Agent is creator of Records resource i inversa.</p>

4.6.2.7.2 Atribut: Rol de l'agent

Identificador	REL007-001
Nom formal	Rol de l'agent
Nom curt	Rol de l'agent
Domini	Defineix el paper que un agent té en relació amb una evidència. Tota evidència té un agent creador (l'agent que ha produït intel·lectualment el document o evidència), i un agent receptor que n'esdevé posteriorment el custodi final, o aquell agent qui finalment l'arxiva.

Tipus de dades	Alfanumèric categòric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	Aquest atribut pot prendre dos valors: <ul style="list-style-type: none"> • Creador • Destinatari
Valor per defecte	No aplica.
Fonts	Correspon a l'element ISAD 3.2.1 Name of creator(s) És equivalent a la relació Unidad documental simple - Agente del model NEDA , on l'agent adquireix el significat de Autor o colaborador en la génesis , i també a la relació més genèrica Unidad documental - Agente , on l'agent adquireix el significat de Destinatario o bé de Remitente . DCAT inclou l'element de <i>Dublin Core</i> dct:creator per indicar el creador d'un recurs.
Exemples	Creador

4.6.2.8 Relació: procés amb procés

4.6.2.8.1 Descripció de la relació

Identificador	REL008
Domini	ENT002 Procés
Rang	ENT002 Procés
Significat de la relació	Relacionat amb
Relació inversa	Relacionat amb
Obligatorietat	No
Cardinalitat	N a N
Definició	Aquesta entitat defineix els dos tipus de relació que poden mantenir dos processos entre ells. Sempre parteix de la relació entre un procés origen i un procés destí, i defineix el rol que el procés origen adopta envers el procés destí.
Atributs	REL008-001 Relació de continuïtat REL008-002 Relació de comprensió
Fonts	Aquesta relació es pot considerar equivalent a l'element 5.3.1 Authorised form(s) of name/Identifier of the related function de la norma ISDF . És equivalent a la relació Función y sus divisiones - Función y sus divisiones del model NEDA .

4.6.2.8.2 Atribut: Relació de continuïtat

Identificador	REL008-001
Nom formal	Relació de continuïtat
Nom curt	Continuïtat
Domini	Aquest atribut defineix el tipus de relació que poden tenir dos processos entre ells situats en una cadena de causa-efecte, i en conseqüència en una cadena temporal.
Tipus de dades	Alfanumèric categòric
Mín. ocurrences	1
Màx. ocurrences	1
Regles de validació	Aquest atribut pot prendre tres valors: <ul style="list-style-type: none">• Antecedent: indica que el procés origen té lloc abans que el procés destí, o com a causa d'aquest, i que la darrera acció o tasca del procés origen determina la primera acció o tasca del procés destí.• Conseqüent: indica que el procés origen té lloc després que el procés destí, o com a conseqüència d'aquest, i que la primera acció o tasca del procés origen ve determinada per la darrera acció o tasca del procés destí.• Depenent: indica que el procés origen té lloc de forma simultània o concurrent amb el procés destí, i que una acció o tasca intermèdia del procés origen determina, o ve determinada per, una acció o tasca del procés destí.
Valor per defecte	No aplica.
Fonts	Aquest atribut es pot considerar equivalent a l'element 5.3.3 Category of relationship de la norma ISDF .
Exemples	Antecedent

4.6.2.8.3 Atribut: Relació de comprensió

Identificador	REL008-002
Nom formal	Relació de comprensió
Nom curt	Comprensió
Domini	Aquest atribut defineix el tipus de relació que poden tenir dos processos entre ells en funció del seu nivell de concreció dins una cadena de jerarquia. Habitualment s'utilitza per agrupar els processos en famílies o macroprocessos.
Tipus de dades	Alfanumèric categòric
Mín. ocurrences	1
Màx. ocurrences	1
Regles de validació	Aquest atribut pot prendre dos valors: <ul style="list-style-type: none">• Conté: indica que el procés origen inclou dins seu el procés destí.

	<ul style="list-style-type: none"> • És part de: indica que el procés origen està inclòs dins el procés destí.
Valor per defecte	No aplica.
Fonts	Aquest atribut es pot considerar equivalent a l'element 5.3.3 Category of relationship de la norma ISDF . I també es podria equiparar a l'element 5.3.2 Type .
Exemples	És part de

4.6.2.9 Relació: procés amb tasca

4.6.2.9.1 Descripció de la relació

Identificador	REL009
Domini	ENT001 Procés
Rang	ENT003 Tasca
Significat de la relació	Es descompon en
Relació inversa	Forma part de
Obligatorietat	Sí
Cardinalitat	1 a N
Definició	Un procés és una seqüència d'activitats o tasques realitzades per agents, en un ordre prèviament conegut o canviant. En el nostre model la tasca, malgrat es realitzi igual en diferents processos, no es considera una entitat independent, sinó una entitat subjecta sempre a un determinat procés. És a través d'aquesta relació que tota tasca es farà dependre d'un procés.
Atributs	Sense atributs propis.
Fonts	<p>Aquesta relació es pot considerar equivalent a l'element 5.3.1 Authorised form(s) of name/Identifier of the related function de la norma ISDF. I també es podria equiparar a l'element 5.3.2 Type.</p> <p>MoReq recull aquesta relació amb la metadada M14.4.65 Participating Class Identifier de l'entitat E14.2.8 Event.</p>

4.6.2.10 Relació: procés amb disposició normativa

4.6.2.10.1 Descripció de la relació

Identificador	REL010
Domini	ENT001 Procés
Rang	ENT010 Disposició normativa
Significat de la relació	Regulat per
Relació inversa	Regula a
Obligatorietat	No
Cardinalitat	N a N
Definició	Tot procés pot estar regulat per una o més disposicions normatives, que en poden regular tant el marc global d'execució com el detall de la seqüència de tasques (el procediment).
Atributs	Sense atributs propis.
Fonts	<p>És equivalent a la relació Normas - Documentos de archivo, Agente y Función y sus divisiones del model NEDA. Aquesta relació es pot considerar equivalent a l'element 5.2.4 Legislation de la norma ISDF.</p> <p>El model MoReq recull aquesta relació en la metadada M14.4.51 Mandate. Però no aplicada al nivell de procés, entitat que no recull explícitament MoReq (ho fa implícitament amb l'entitat E14.2.2 Class), sinó al nivell de l'entitat E14.2.6 Disposal Schedule.</p> <p>Correspon a la relació RiC-R063 Thing is regulated by rule i inversa.</p>

4.6.2.11 Relació: tasca amb tasca

4.6.2.11.1 Descripció de la relació

Identificador	REL011
Domini	ENT003
Rang	ENT003
Significat de la relació	Antecedent de
Relació inversa	Conseqüent a
Obligatorietat	Sí
Cardinalitat	N a N
Definició	Tota tasca té una o varies tasques anteriors, que condicionen la seva execució, llevat de les tasques que inicien un procés. De la mateixa manera, tota tasca té una o varies tasques conseqüents,

	llevat de les tasques que finalitzen el procés. En el cas dels <i>gateways</i> , les tasques conseqüents no depenen de l'execució de la tasca anterior, sinó de les condicions que s'hagin de complir i que determinen la bifurcació del flux.
Atributs	REL011-001 Rol de la tasca
Fonts	<p>Aquesta relació es pot considerar equivalent a l'element 5.3.1 Authorised form(s) of name/Identifier of the related function de la norma ISDF.</p> <p>MoReq recull aquesta relació amb la metadada M14.4.71 Participating Event Identifier de l'entitat E14.2.8 Event.</p> <p>RiC contempla la relació RiC-R006 Event is a subevent of Event i inversa, com a relació de comprensió, però no com a relació de conseqüència. Aquesta relació la trobem expressada en la relació RiC-R008 Thing precedes/follows Thing (directa/inversa).</p>

4.6.2.11.2 Atribut: Rol de la tasca

Identificador	REL011-001
Nom formal	Rol de la tasca
Nom curt	Rol de la tasca
Domini	La relació entre dues tasques es caracteritza per l'ordre d'execució del flux de tasques, que assigna el rol d'antecedent o conseqüent.
Tipus de dades	Alfanumèric categòric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	<p>L'atribut pot prendre dos valors:</p> <ul style="list-style-type: none"> • Antecedent • Conseqüent
Valor per defecte	No aplica.
Fonts	Aquest atribut es pot considerar equivalent a l'element 5.3.3 Category of relationship de la norma ISDF .
Exemples	Antecedent

4.6.2.12 Relació: sèrie documental amb sèrie documental

4.6.2.12.1 Descripció de l'entitat

Identificador	REL012
Domini	ENT007 Sèrie documental
Rang	ENT007 Sèrie documental

Significat de la relació	Relacionada amb
Relació inversa	Relacionada amb
Obligatorietat	No
Cardinalitat	N a N
Definició	Aquesta entitat defineix el tipus de relació que poden mantenir dues sèries entre elles. Sempre parteix de la relació entre una sèrie origen i una sèrie destí, i defineix el rol que la sèrie origen adopta envers la sèrie destí.
Atributs	REL012-001 Rol de la sèrie
Fonts	La norma ISAD contempla la relació entre sèries documentals en l'element 3.5.3 Related units of description . EAD disposa de dos elements similars, també amb possible equivalència amb aquesta relació: <relatedmaterial> i <separatedmaterial> . És equivalent a la relació Documentos de archivo – Documentos de Archivo del model NEDA .

4.6.2.12.2 Atribut: Rol de la sèrie

Identificador	REL012-001
Nom formal	Rol de la sèrie
Nom curt	Rol de la sèrie
Domini	Aquest atribut defineix el tipus de relació que poden tenir dues sèries situades en una cadena de causa-efecte, i en conseqüència en una cadena temporal. A diferència dels processos, en aquest cas en vincle de relació entre les sèries no són les accions pures sinó determinades evidències resultants de les accions.
Tipus de dades	Alfanumèric categòric
Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	Aquest atribut pot prendre quatre valors: <ul style="list-style-type: none"> • Antecedent: indica que la sèrie origen completa la seva formació abans que s'iniciï la formació de la sèrie destí, i que el contingut informatiu final de la sèrie origen (la matriu de variables) determina el contingut informatiu de les primeres evidències de la sèrie destí. Aquesta relació requereix sempre que els processos generadors de les sèries, en cas de ser diferents, mantinguin també una relació de continuïtat en el mateix sentit que les sèries. • Conseqüent: indica que la sèrie origen inicia la seva formació un cop s'ha completat la formació de la sèrie destí, i que el contingut informatiu final de la sèrie destí (la matriu de variables) determina el contingut informatiu de les primeres evidències de la sèrie origen. Aquesta relació requereix sempre que els processos generadors de les

	<p>sèries, en cas de ser diferents, mantinguin també una relació de continuïtat en el mateix sentit que les sèries.</p> <ul style="list-style-type: none"> • Relacionada: indica que la formació de la sèrie origen i destí coincideixen en el temps, i que determinades evidències intermèdies de la sèrie origen poden existir també en la sèrie destí. Aquesta relació pot prendre dues dimensions:²²² <ul style="list-style-type: none"> ○ Si les dues sèries són resultat del mateix procés, però tenen diferent productor, aquesta relació s'anomenarà una relació de derivació. En funció del rol del productor dins el procés, del caràcter de les evidències, i del grau de completesa en la forma com la sèrie documenta totes les accions o tasques del procés generador, una sèrie podrà adquirir el rol de sèrie primària, i les altres sèries derivades adquiriran el rol de sèries secundàries. ○ Si les dues sèries són resultat de processos diferents, cap de les dues sèries adquirirà cap rol específic, més enllà de la constància de l'existència d'una relació. • Recapitulativa: indica que la formació de la sèrie origen és conceptualment posterior a la sèrie destí, malgrat que temporalment puguin ser simultànies en la seva formació. La sèrie origen adquireix la seva condició de sèrie recapitulativa en el moment en què recull una part succinta de les dades d'una o més evidències de la sèrie destí.
Valor per defecte	No aplica.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	Antecedent

4.6.2.13 Relació: procés amb sèrie documental

4.6.2.13.1 Descripció de la relació

Identificador	REL013
Domini	ENT002 Procés
Rang	ENT007 Sèrie documental
Significat de la relació	Dona origen a
Relació inversa	És resultat de
Obligatorietat	Sí
Cardinalitat	1 a N

²²² Les dues dimensions que pot prendre aquesta relació es recullen en l'atribut **ENT007-004 Caràcter** (de la sèrie documental). Per a més informació sobre la seva significació, vegeu l'apartat 2.3.1.4.2 de l'estat de la qüestió (capítol 2).

Definició	Tot procés, en executar-se, genera evidències, que s'agrupen en forma de sèries documentals. Cada procés que s'hagi executat haurà generat com a mínim una sèrie documental. De la mateixa manera, tota sèrie documental és sempre el resultat d'un sol procés. El procés determina, doncs, els límits funcionals de la sèrie documental.
Atributs	Sense atributs propis.
Fonts	<p>És equivalent a la relació Documentos de archivo - Función y sus divisiones del model NEDA.</p> <p>Equival a l'atribut classification que preveu la norma ISO 23081.</p> <p>Correspon a la relació RiC-R033 Activity is documented by Record resource i inversa.</p>

4.6.2.14 Relació: sèrie documental amb regla d'accés

4.6.2.14.1 Descripció de la relació

Identificador	REL014
Domini	ENT007 Sèrie documental
Rang	ENT008 Regla d'accés
Significat de la relació	Regulada per
Relació inversa	Regula l'accés a
Obligatorietat	Sí
Cardinalitat	1 a N
Definició	Tota sèrie documental té com a mínim una regla que regula l'accés d'un agent a les evidències de la sèrie. La regla d'accés regula els privilegis amb els quals un agent accedeix a les evidències d'una sèrie en un moment o termini determinats. En cas que més agents hagin d'accedir a les evidències de la sèrie, o bé ho facin en diferents moments, es definirà en cada cas una nova regla d'accés.
Atributs	Sense atributs propis.
Fonts	<p>És equivalent a la relació Normas - Documentos de archivo, Agente y Función y sus divisiones del model NEDA.</p> <p>Equival a l'atribut access que preveu la norma ISO 23081.</p> <p>Correspon a la relació RiC-R063 Thing is regulated by rule i inversa.</p>

4.6.2.15 Relació: sèrie documental amb regla de conservació

4.6.2.15.1 Descripció de la relació

Identificador	REL015
Domini	ENT007 Sèrie documental
Rang	ENT009 Regla de conservació
Significat de la relació	Regulada per
Relació inversa	Regula la conservació de
Obligatorietat	Sí
Cardinalitat	1 a N
Definició	Tota sèrie documental té com a mínim una regla de conservació. Per una mateixa sèrie es crearan tantes regles de conservació com resolucions i terminis s'estableixin en el cicle de vida de les evidències d'aquella sèrie.
Atributs	Sense atributs propis.
Fonts	<p>MoReq materialitza l'assignació de regla de conservació a través de la metadada M14.4.11 Default Disposal Schedule Identifier. Però ho fa al nivell de l'entitat E14.2.2 Class, que equipara a classe de d'un quadre de classificació funcional, i no al nivell de sèrie documental. MoReq planteja un model on les agrupacions documentals hereten el cicle de vida directament d'una entrada del quadre de classificació, sense incloure explícitament el concepte de sèrie (E14.2.1 Aggregation) en cap de les dues entitats.</p> <p>És equivalent a la relació Normas - Documentos de archivo, Agente y Función y sus divisiones del model NEDA.</p> <p>Correspon a la relació RiC-R063 Thing is regulated by rule i inversa.</p>

4.6.2.16 Relació: procés amb evidència

4.6.2.16.1 Descripció de la relació

Identificador	REL016
Domini	ENT002 Procés
Rang	ENT004 Evidència
Significat de la relació	Genera
Relació inversa	Generada per
Obligatorietat	Sí
Cardinalitat	1 a N

Definició	L'execució d'un procés genera evidències. Cada evidència és el producte d'un procés i té significació en el marc d'aquell procés.
Atributs	Sense atributs propis.
Fonts	És equivalent a la relació Documentos de archivo - Función y sus divisiones del model NEDA . Correspon a la relació RiC-R033 Activity is documented by Record resource i inversa.

4.6.2.17 Relació: tasca amb evidència

4.6.2.17.1 Descripció de la relació

Identificador	REL017
Domini	ENT003 Tasca
Rang	ENT004 Evidència
Significat de la relació	Genera/rep
Relació inversa	Generada/rebuda per
Obligatorietat	No/Sí
Cardinalitat	N a N
Definició	Tota tasca necessita informació per poder-se executar, i genera a la seva vegada nova informació. Aquesta informació, tant d'entrada com de sortida, es manifesta en forma d'evidències. Per aquest motiu, tota tasca pot rebre diverses evidències d'entrada, i generar diverses evidències de sortida.
Atributs	REL017-001 Sentit de l'evidència
Fonts	MoReq recull aquesta relació amb la metadada M14.4.77 Participating Record Identifier de l'entitat E14.2.8 Event . Correspon a la relació RiC-R033 Activity is documented by Record resource i inversa.

4.6.2.17.2 Atribut: Sentit de l'evidència

Identificador	REL017-001
Nom formal	Sentit de l'evidència
Nom curt	Sentit de l'evidència
Domini	L'atribut indica el sentit de l'evidència en relació amb una tasca. Si es rep l'evidència per consumir la seva informació s'anomenarà evidència d'entrada. Si, per contra, es genera per contenir informació generada per la tasca, s'anomenarà evidència de sortida.
Tipus de dades	Alfanumèric categòric

Mín. ocurrencies	1
Màx. ocurrencies	1
Regles de validació	L'atribut pot prendre dos valors: <ul style="list-style-type: none"> • Entrada • Sortida
Valor per defecte	No aplica.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	Entrada

4.6.2.18 Relació: sèrie documental amb evidència

4.6.2.18.1 Descripció de la relació

Identificador	REL018
Domini	ENT007 Sèrie documental
Rang	ENT004 Evidència
Significat de la relació	Conté
Relació inversa	Continguda en
Obligatorietat	Sí
Cardinalitat	1 a N
Definició	La sèrie documental es compon del conjunt d'evidències acumulades pel productor de la sèrie com a resultat de la seva participació en un procés. Considerarem que una evidència és acumulada pel productor quan l'agent que té el rol de productor en relació amb la sèrie documental apareix com a destinatari o receptor de l'evidència.
Atributs	Sense atributs propis.
Fonts	El model AGRkMS disposa de la propietat 1 Category que, aplicada a l'entitat Record , serveix per relacionar-la amb la sèrie documental de la qual forma part. Equival a l'atribut classification que preveu la norma ISO 23081 . És equivalent a la relació Documentos de archivo – Documentos de archivo del model NEDA (relació jeràrquica). Correspon a la relació RiC-R024 Record set includes record i inversa .

4.6.2.19 Relació: sistema d'informació amb evidència

4.6.2.19.1 Descripció de la relació

Identificador	REL019
Domini	ENT011 Sistema d'informació
Rang	ENT004 Evidència
Significat de la relació	Conté
Relació inversa	Continguda en
Obligatorietat	No/Sí
Cardinalitat	1 a N
Definició	Aquesta relació s'aplica només per a les evidències del tipus "Dades" (atribut ENT004-005). Les evidències de tipus dades corresponen a interfícies de visualització de dades (pantalles d'entrada o sortida de dades) que un usuari o aplicació utilitza en un moment determinat de l'execució d'un procés. Tota interfície de visualització pertany a un sistema d'informació.
Atributs	Sense atributs propis.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.

4.6.2.20 Relació: evidència amb dada elemental

4.6.2.20.1 Descripció de l'entitat

Identificador	REL020
Domini	ENT004 Evidència
Rang	ENT006 Dada elemental
Significat de la relació	Conté
Relació inversa	Apareix a
Obligatorietat	Sí
Cardinalitat	N a N
Definició	Tota evidència conté dades, en forma de variables d'informació que poden contenir valors o requerir l'aportació d'un valor per part d'un agent. Aquesta relació defineix la forma com una dada apareix dins una evidència, i quin valor pren com a resultat de l'execució d'una tasca d'un procés. En el context d'aquest model aquesta relació pot aparèixer descrita amb el terme "ocurrència" o "instanciació" d'una dada elemental dins una evidència.
Atributs	REL020-001 Tipus d'ocurrència
Fonts	Es pot considerar equivalent a la relació Documentos de archivo – Documentos de archivo, Agente, Función y sus divisiones,

	Norma, Concepto, Objeto o Acontecimiento y Lugar del model NEDA.
--	---

4.6.2.20.2 Atribut: Tipus d'ocurrència

Identificador	REL020-001
Nom formal	Tipus d'ocurrència
Nom curt	Tipus d'ocurrència
Domini	Indica si la variable d'informació de l'evidència ja conté informació en el moment de creació de l'evidència, o si aquesta informació (el valor d'una dada) es genera durant el procés de creació de l'evidència.
Tipus de dades	Alfanumèric categòric
Mín. ocurrences	1
Màx. ocurrences	1
Regles de validació	<p>Aquest atribut pot prendre dos valors:</p> <ul style="list-style-type: none"> • Autòctona: quan la variable d'informació no existeixi en una forma informada en el moment en què s'inicia la creació de l'evidència, i s'empleni amb un valor com a resultat de l'acció o tasca que genera aquella evidència. Aquestes variables es poden emplenar amb una acció manual per part d'un usuari o bé de forma automàtica a partir d'algun tipus de càlcul, fórmula o operació realitzada dins la pròpia evidència. • Importada: quan la variable d'informació ja contingui un valor quan s'inicia la creació de l'evidència. El terme importada descriu el fet que el valor de la variable no es genera durant la creació de l'evidència, sinó que s'ha generat en una acció o tasca anterior, i pot aparèixer preempletat per facilitar informació de suport al creador de l'evidència. Les variables importades d'una evidència són variables que ja han aparegut en evidències anteriors del procés o, en el cas de la primera evidència del procés, que s'han creat en evidències de processos anteriors.
Valor per defecte	"Autòctona"
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
Exemples	Autòctona

4.6.2.21 Relació: evidència amb model de document

4.6.2.21.1 Descripció de la relació

Identificador	REL021
Domini	ENT004 Evidència
Rang	ENT005 Model de document
Significat de la relació	Es formalitza amb
Relació inversa	Formalitza la forma externa de
Obligatorietat	No/Sí
Cardinalitat	N a N
Definició	Una evidència pot utilitzar una plantilla o model de document com a recurs per formalitzar la seva aparença, forma externa, o funcionament pel que fa a l'entrada de dades. Atès que l'evidència és un concepte, poden existir més d'un model aplicables a una mateixa evidència d'un mateix procés, o bé evidències on la formalització sigui opcional (per exemple en processos de relació amb la ciutadania on la formalització no sigui un requisit obligatori).
Atributs	Sense atributs propis.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.

4.6.2.22 Relació: interfície amb dada elemental

4.6.2.22.1 Descripció de la relació

Identificador	REL022
Domini	ENT012 Interfície
Rang	ENT006 Dada elemental
Significat de la relació	Trasllada
Relació inversa	Traslladada/migrada per
Obligatorietat	No
Cardinalitat	N a N
Definició	La interfície o connexió entre sistemes es fa bàsicament per a intercanviar dades entre els sistemes, i de forma especial en aquest model, on l'entitat sistema d'informació sempre es vincula a un sistema contenidor de dades. Per aquest motiu, aquesta connexió s'anomena sovint flux de dades (<i>dataflow</i>). Aquesta relació indica quines dades elementals són traslladades amb cada interacció o connexió entre sistemes d'informació, i reben el nom d'ocurrències de la dada elemental en la interfície o flux de dades.

Atributs	Sense atributs propis.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.

4.6.2.23 Relació: sistema d'informació amb dada elemental

4.6.2.23.1 Descripció de la relació

Identificador	REL023
Domini	ENT011 Sistema d'informació
Rang	ENT006 Dada elemental
Significat de la relació	Conté
Relació inversa	Continguda a
Obligatorietat	Sí/No
Cardinalitat	N a N
Definició	Tot sistema d'informació emmagatzema dades (o variables d'informació). Aquest conjunt, un cop estructurat, configura el model de dades que consumeix el sistema d'informació quan executa les seves funcionalitats. Aquesta relació descriu doncs l'aparició de dades elementals en el model de dades d'un determinat sistema d'informació, relació per a la qual també s'utilitzarà el terme "ocurrència" (de la dada elemental en el sistema).
Atributs	Sense atributs propis.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.

4.6.2.24 Relació: sistema d'informació amb objecte

4.6.2.24.1 Descripció de la relació

Identificador	REL024
Domini	ENT011 Sistema d'informació
Rang	ENT014 Objecte
Significat de la relació	Descriu de forma principal
Relació inversa	És unitat de descripció de
Obligatorietat	Sí/No
Cardinalitat	N a 1

Definició	<p>El sistema d'informació, des del punt de vista conceptual, acumula dades sobre un objecte, que esdevé la seva unitat de descripció principal. Aquesta atribució es fa habitualment de forma intuïtiva, i es recull en els requeriments funcionals que posteriorment determinen el model de dades. Això no obstant, a través del model de dades es pot conèixer el mapa complet d'objectes descrits per les dades del sistema d'informació. Aquesta relació recull, doncs, l'objecte que s'ha considerat principal durant el disseny del sistema pel que fa a la definició del model de dades.</p> <p>El concepte d'unitat de descripció és un concepte propi de la normativa de descripció arxivística, que significa aquella cosa del món real que és objecte de descripció (un expedient, un document, etc.). És equivalent al concepte d'objecte de la dada utilitzat en la relació REL026 d'aquest model.</p>
Atributs	Sense atributs propis.
Fonts	<p>Malgrat RiC no contempla l'entitat sistema d'informació en la forma que es fa en aquest model, si s'equipara a sèrie documental s'hi pot aplicar l'equivalència amb la relació RiC-R020 Record resource has main subject Thing i inversa.</p> <p>L'element més similar a aquest del model DCAT és l'element dcat:theme. Un recurs descrit amb DCAT pot tenir múltiples temes.</p>

4.6.2.25 Relació: objecte amb objecte

4.6.2.25.1 Descripció de la relació

Identificador	REL025
Domini	ENT014 Objecte
Rang	ENT014 Objecte
Significat de la relació	Subtipus de
Relació inversa	Supertipus de
Obligatorietat	No
Cardinalitat	N a 1
Definició	<p>Es contempla en aquest model una relació de comprensió entre objectes. Malgrat que aquesta relació pot tenir un desenvolupament ontològic molt més complet, per motius de practicitat de cara a l'explotació de la informació recollida, en aquest model s'ha limitat aquesta relació a dos nivells, i s'ha utilitzat per descriure el nivell superior el concepte de superobjecte.</p>
Atributs	Sense atributs propis.

Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.
-------	--

4.6.2.26 Relació: objecte amb dada elemental

4.6.2.26.1 Descripció de la relació

Identificador	REL026
Domini	ENT014 Objecte
Rang	ENT006 Dada elemental
Significat de la relació	Descrit per
Relació inversa	Descriu
Obligatorietat	Sí
Cardinalitat	1 a N
Definició	Cada objecte, sigui cosa o acció, és descrit per un conjunt de dades elementals. Aquestes dades descriuen només aquell objecte, malgrat el nom i el tipus de la dada siguin el mateix per a diferents objectes. L'objecte esdevé així la forma bàsica de relacionar les dades entre elles. Aquesta relació reflecteix la dependència de tota dada elemental de l'objecte al qual descriu.
Atributs	Sense atributs propis.
Fonts	Es pot considerar equivalent a la relació Documentos de archivo – Documentos de archivo, Agente, Función y sus divisiones, Norma, Concepto, Objeto o Acontecimiento y Lugar del model NEDA . Malgrat RiC no contempla l'entitat dada elemental, podem trobar punts de contacte amb la relació RiC-R019 Record resource has subject Thing i inversa, i també amb RiC-R021 Record resource describes Thing i inversa.

4.6.2.27 Relació: Visualització de dades amb dada elemental

4.6.2.27.1 Descripció de la relació

Identificador	REL027
Domini	ENT013 Visualització de dades
Rang	ENT006 Dada elemental
Significat de la relació	Conté
Relació inversa	Apareix a

Obligatorietat	Sí/No
Cardinalitat	N a N
Definició	Tota visualització de dades utilitza dades elementals. La dada elemental representada en una visualització serà la mateixa amb independència del tipus de component visual que s'utilitzi. Novament l'aparició d'una dada elemental en una visualització de dades la definirem amb el terme "ocurrència".
Atributs	Sense atributs propis.
Fonts	No s'ha localitzat en els models analitzats cap element que es pugui considerar equivalent a l'element que es descriu en aquest apartat.

4.6.2.28 Relació: sistema d'informació amb sistema d'informació (interfície)

En aquest model s'ha considerat que la relació entre sistemes d'informació sempre requereix, si és manual, un conjunt d'accions de trasllat de dades, i si és automàtica algun tipus de component que faci efectiu aquest trasllat. Aquesta relació ha donat origen, doncs, a l'entitat anomenada "Interfície", que s'ha identificat amb el codi **ENT012**, i que permet recollir les relacions d'intercanvi de dades entre sistemes d'informació i aplicacions. Vegeu la descripció d'aquesta entitat en l'apartat 4.6.1.12 d'aquest capítol.

4.6.3 Correspondència de les entitats amb cada estàndard

Taula 4-2: Correspondència terminològica entre les entitats del model de dades i els models de referència

	ENT001 Agent	ENT002 Procés	ENT003 Tasca	ENT004 Evidència	ENT006 Dada	ENT007 Sèrie	ENT010 Normativa	ENT011 Sistema	ENT014 Objecte
AGRkMS	Agent	Business		Record		Record	Mandate	System (agent)	
DCAT						Resource		Distribution (dataset)	Concept (SKOS)
EAC	Entity or person								
EAD				Item		Series			
e-EMGDE	Agente	Actividad		Documento		Serie	Regulación	Dispositivo (agente)	
ISAAR	Corporate body, person or family	Function							
ISAD				Item		Series			
ISDF		Function	Task						
ISDIAH									
ISO 23081	People (agent)	Business		Record		Record	Mandate		
MoReq	User	Class	Event	Record		Aggregation			
NEDA	Agente	Función	Acción	Unidad documental	Componente documental	Serie	Norma		Concepto, objeto, acontecimiento
ONDC				Information asset		Information asset	Legal authority	Information asset	
PROV	Agent	Activity							Entity / Activity
RiC	Agent	Activity		Record	Record part	Record set	Rule / Mandate	Mechanism (agent)	Thing
SIARD					Column/field	View		Database	

Les entitats **ENT008 Regla d'accés**, **ENT009 Regla de conservació** i **ENT012 Interfície de connexió entre sistemes** s'expressen en pràcticament tots els models com atributs, no com entitats. Per aquest motiu no s'han inclòs en aquest quadre.

De les entitats **ENT005 Model de document** i **ENT013 Visualització de dades** no s'han trobat mencions directes a les normes de descripció analitzades. En ambdós casos, si se'n fa esment, es relacionen amb atributs del concepte de document o *record* (en el cas del model de document) i del concepte de conjunt de dades o *dataset* (en el cas de la visualització de dades). Per aquest motiu tampoc s'han inclòs en aquest quadre.

Capítol 5. Proposta d'indicadors per a la governança de la informació

5.1 Plantejament metodològic

Un dels elements cabdals de qualsevol sistema de governança és la capacitat de mesurar magnituds de tots els elements del sistema que es vol governar, i posar-les en relació (Kampffmeyer, 2023). Aquesta condició, aplicada a la governança de la informació, més enllà dels usos preceptius en el marc de la gestió de riscos i el *compliance*, és la clau per la millora efectiva dels processos i la informació d'una organització (Khatri & Brown, 2010; Kooper et al., 2011; Ojala, 2021). Això inclou l'ús de solucions d'anàlisi estadística i d'eines de visualització de dades (Schmid et al., 2019; Smajevic & Bork, 2021).

Un dels vuit principis GARP de *ARMA Internacional*²²³ és el nivell de *compliance*, i en les accions que estableix per assolir el nivell 5 del model de maduresa GARP (maduresa transformacional) s'indica que "*Auditing and continuous improvement processes are well established and monitored by senior management*". L'objectiu d'aquesta monitorització és disposar d'un coneixement detallat, i alhora exhaustiu i global, de l'estat de cada element en cada moment, i de l'impacte de cada canvi que s'hi produeixi, sigui aquest canvi causat per fets externs o per les pròpies accions de governança, en la línia de la visió procedimental de la governança propugnada per Loshin i Ladley (Loshin, 2010; Ladley, 2016). Però totes les experiències apunten a que no és una tasca fàcil (Benfeldt et al., 2020; Hoke, 2022; Torreblanca, 2023).

Les accions de governança tenen bàsicament dos objectius:

1. El primer és poder optimitzar el funcionament del sistema que es governa, poder saber què cal canviar per aconseguir un determinat resultat.
2. I el segon és protegir el sistema dels riscos, tant interns com externs, que puguin perjudicar el seu funcionament.

Ulrich Kampffmeyer distingeix la governança de la mera gestió amb aquesta asserció (Kampffmeyer, 2023):

²²³ *Generally Accepted Recordkeeping Principles* (GARP), de *ARMA Internacional*, disponibles a <https://www.arma.org/page/principles>. Per al model de maduresa vegeu <https://www.arma.org/page/PrinciplesMaturityModel>.

“Management is doing what we have to do. Governance is doing what we want to do.”

Per tant, la base d'una bona governança és la capacitat de **control**, que és el que proporciona la suma dels dos objectius anteriorment esmentats.

Als efectes d'assolir aquesta capacitat de control en la governança d'un sistema de gestió documental orientat a dades (dadificat), es proposen en aquest capítol un conjunt de 83 indicadors de mesura. Aquests indicadors s'han identificat i construït a partir del treball realitzat sobre el cas objecte d'estudi, el Consorci d'Educació de Barcelona, la metodologia exposada a la **GMAP** i la **GDAD**, les aportacions de diferents autors (Caplinger & Borgschulte, 2015; Penteado, 2018), i les necessitats de governança observades ja durant l'execució del propi projecte.²²⁴

Malgrat aquesta vinculació a una organització en concret, s'ha intentat construir els indicadors amb una vocació prou genèrica com per a que puguin ser aplicables a d'altres organitzacions de característiques i dimensions similars. El Consorci d'Educació de Barcelona, l'ens on s'han posat a prova aquests indicadors, és un ens quantitativament i qualitativa representatiu de bona part del sector públic català i de l'estat espanyol, tant pel que fa a l'abast competencial (i per tant nombre d'ítems conceptuals a gestionar) com a la dinàmica de treball, estructura organitzativa i tecnologia utilitzada.²²⁵ Per aquest

²²⁴ El caràcter canviant de l'organització objecte d'estudi, que ha estat modificant constantment la informació que s'anava recollint, ha posat de relleu la necessitat de disposar des de l'inici del projecte d'uns indicadors bàsics que evitin la pèrdua de control sobre l'evolució del sistema que es construeix.

²²⁵ D'acord amb el *Banc de dades d'ocupació pública*, disponible al *Portal de dades obertes* de la Generalitat de Catalunya (<https://analisi.transparenciacatalunya.cat/Sector-P-blic/Banc-de-dades-d-ocupaci-p-blica-Treballadors-de-l-/n4nv-s873>), el nombre de treballadors dels consorcis i ens autònoms de titularitat pública no territorialitzats (els que tenen competències transversals i seu a Barcelona), considerant el personal estructural i conjuntural, oscil·la entre 50 i 300 treballadors, llevat d'algunes excepcions (com el Consorci de Normalització Lingüística). En aquest sentit, el Consorci d'Educació de Barcelona se situa en la franja alta, amb 236 treballadors/ores a la seu central, i prop d'un centenar més als serveis educatius de districte. En el cas de l'Ajuntament de Barcelona, i segons les dades del Repositori Obert de Coneixement (BCNROC) de l'Ajuntament de Barcelona, disponible a <https://bcnroc.ajuntament.barcelona.cat/jspui/handle/11703/127498>, les magnituds de la plantilla del sector públic municipal són similars (entre 50 i 300 en seu central).

En aquesta magnitud no s'inclou el personal docent i PAS que conforma la plantilla dels 316 centres educatius públics de totes les tipologies i nivells educatius que gestiona directament el Consorci d'Educació de Barcelona, atès que no formen part directament de la plantilla del Consorci, sinó del Departament d'Educació i de l'Ajuntament de Barcelona. D'acord amb el principi d'autonomia de centre, regulat pel Decret 102/2010, de 3 d'agost, d'autonomia dels centres educatius, cada centre gestiona els seus recursos de forma autònoma. Per aquest motiu, els 217 processos que gestiona el Consorci són atesos exclusivament pels 236 treballadors del Consorci, juntament amb el personal dels serveis educatius.

Pel que fa a l'arquitectura tecnològica, el Consorci està integrat en el model de governança TIC del Centre de Telecomunicacions i Tecnologies de la Informació de la Generalitat de Catalunya (CTTI), per la qual cosa la tecnologia i recursos que té al seu abast són exactament els mateixos

motiu, es considera que aquests indicadors poden ser aplicables a d'altres ens de característiques similars.

Els indicadors són mesures.²²⁶ Aquestes mesures es poden representar amb diferents nivells d'agregació, i amb determinats filtres de valors, que s'indiquen en la descripció del propi indicador. Això no obstant, sempre es podran aplicar d'altres nivells d'agregació als previstos en aquesta descripció. Cada indicador està identificat amb un codi, que està compost d'un prefix amb les inicials de l'entitat analitzada (PRO, SER, SIS o DAD) i d'un número continu dins aquella entitat, separats per un guió.

En total s'ha construït tres tipus d'indicadors numèrics (A, B i C), que es complementen amb dues eines d'anàlisi visuals (D i E), els quals s'han agrupat en funció de les quatre entitats principals del model de dades: els processos, les sèries documentals, els sistemes d'informació, i les dades elementals. Els tipus d'indicadors i eines d'anàlisi són els següents:

A. Mètriques descriptives basades únicament en atributs de l'entitat analitzada. Cada mètrica es descriu amb el requadre següent:

[CODI DE LA MÈTRICA]	
Indicador global	Nombre de [ENTITAT ANALITZADA]
Nivells d'agregació	[NOM DELS ATRIBUTS DE L'ENTITAT ANALITZADA PELS QUALS ES PROPOSA AGREGAR L'INDICADOR]

B. Mètriques descriptives derivades de les relacions de l'entitat analitzada amb altres entitats. Aquestes mètriques mesuren les correspondències entre entitats i proposen agregacions en funció dels valors d'atributs de les entitats relacionades. Cada mètrica es descriu amb el requadre següent:

[CODI DE LA MÈTRICA]	
Indicador de [ENTITAT ANALITZADA]	Nombre de [ENTITAT RELACIONADA]
Nivells d'agregació	[NOM DELS ATRIBUTS DE L'ENTITAT RELACIONADA PELS QUALS ES PROPOSA AGREGAR L'INDICADOR]

C. Indicadors d'anàlisi. Aquests indicadors es basen en càlculs i combinacions realitzats a partir dels dos tipus de mètriques descriptives anteriors, i es presenten

que qualsevol altre ens de la Generalitat de Catalunya i del seu sector públic, i molt similars als de la resta d'Administracions catalanes i estatals.

²²⁶ Utilitzen en els encapçalaments d'apartat el terme "mètrica" per fer referència als valors que pot prendre cada indicador descrit en aquest document.

en forma d'índex o, si escau, de percentatges. Cada indicador es descriu amb el requadre següent:

[CODI DE L'INDICADOR]		
Indicador de [ENTITAT ANALITZADA]	Nom de l'indicador	Interpretació de la mesura
Indicadors de base: [CODI I NOM DELS INDICADORS DESCRIPTIUS UTILITZATS PER AL CàLCUL]	[FÓRMULA DE CàLCUL DE L'INDICADOR] ²²⁷	[DESCRIPCIÓ DEL SIGNIFICAT DE DETERMINATS VALORS DE L'INDICADOR] ²²⁸

D. **Matriu de variables.** Aquesta és la visualització per defecte per mostrar l'evolució de les dades durant el desenvolupament d'un procés o d'una sèrie documental. Només s'aplica a aquestes dues entitats. Això no exclou que el mateix patró de visualització es pugui utilitzar quan es comparen relacions estàtiques de variables (diccionaris de variables de processos, sèries o sistemes) de diferents entitats consecutives o relacionades. Però l'entorn on s'aplicarà habitualment serà per mostrar l'evolució de les variables en un circuit documental.²²⁹ L'estructura que adopta en aquest cas és la següent:

	[Evidència 1]	[Evidència 2]	[Evidència 3]	[Evidència N]
[Dada elemental 1]	[Ocurrència 1]	[Ocurrència 2]		[Ocurrència 3]
[Dada elemental 2]	[Ocurrència 1]	[Ocurrència 2]	[Ocurrència 3]	
[Dada elemental 3]		[Ocurrència 1]		[Ocurrència 2]
[Dada elemental N]				[Ocurrència 1]

E. **Mapa de relacions.** Totes les entitats que es relacionen es poden representar en forma de mapa de relacions, on els nodes són cadascuna de les entitats que estan relacionades, i els vectors cadascuna de les relacions que mantenen entre elles. Els mapes permeten una anàlisi visual, i anàlisi focals de determinats conjunts de relacions. Els mapes de referència són els que representen relacions entre un mateix tipus d'entitat, com per exemple mapa de processos relacionats, de sèries relacionades o d'interaccions entre sistemes d'informació.

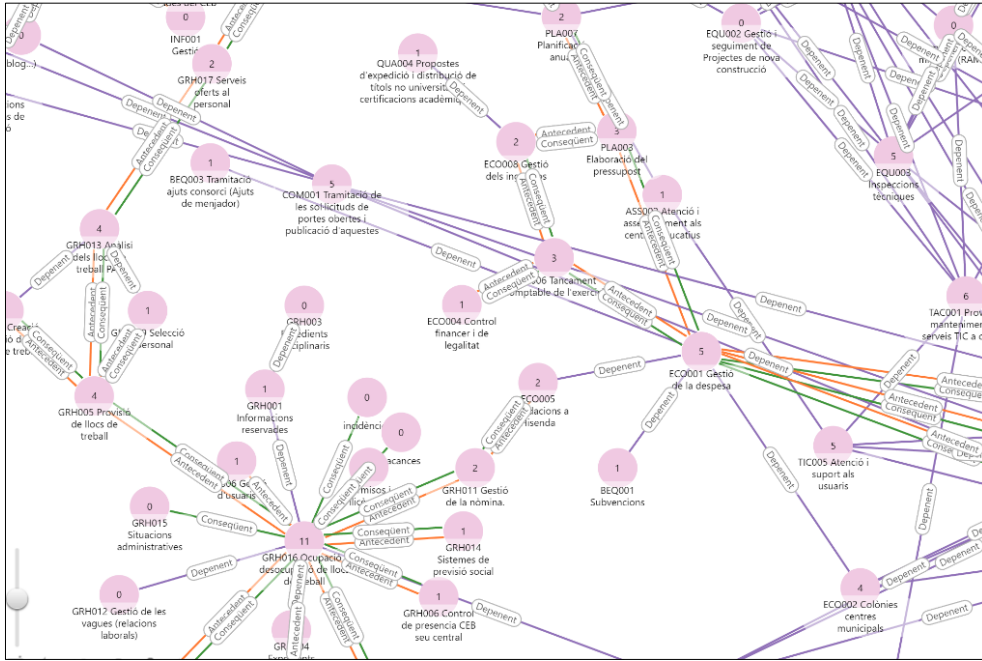
²²⁷ Per a la representació de les fórmules de càlcul s'ha experimentat amb una proposta de notació, de desenvolupament propi, que s'ha buscat que fos el més intuïtiva possible. Aquesta proposta utilitza els següents símbols (a part dels operadors lògics convencionals, AND i OR):

- []: variable de càlcul
- (): precedència de les operacions
- Flt: aplicació d'un filtre o condició
- “ “: valor de condició

²²⁸ La interpretació que es proposa en els indicadors d'anàlisi és producte d'una inferència, que en alguns casos no ha pogut ser demostrada a partir del cas d'aplicació.

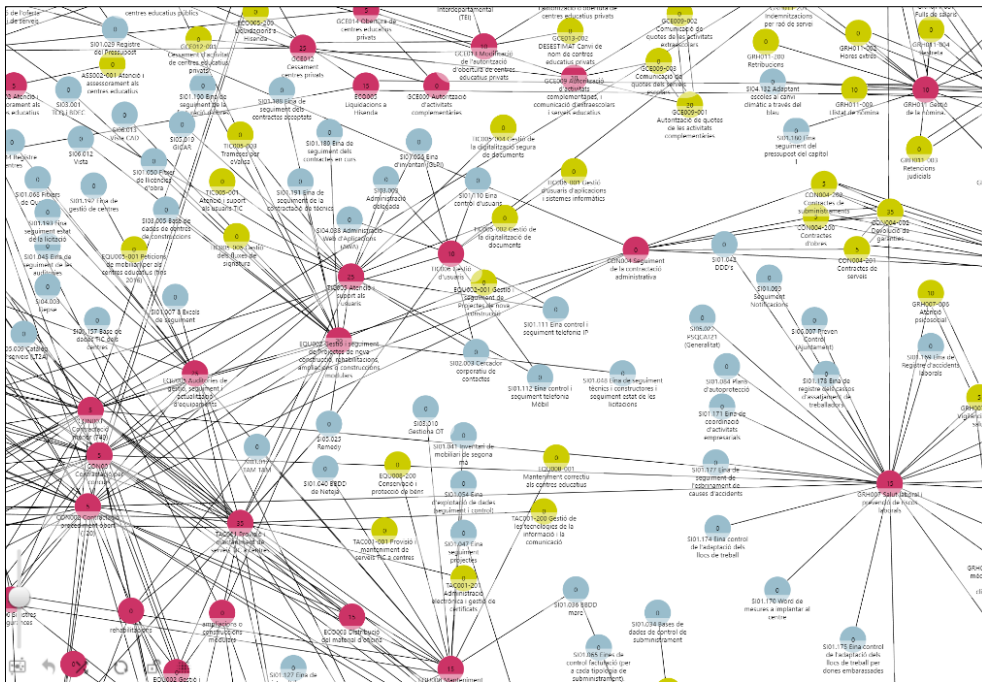
²²⁹ En el context d'aquesta tesi, només es mostren les matrius de variables de la mostra de processos dadificats de l'apartat 7.1 del capítol 7.

Figura 5-1: Exemple del mapa de processos del Consorci d'Educació de Barcelona



Però també inclou els mapes que combinen diferents tipus d'entitats relacionades entre elles.

Figura 5-2: Exemple del mapa global d'actius d'informació del Consorci d'Educació de Barcelona



A continuació s'enumeren i es descriuen breument cadascun els indicadors, agrupats per l'entitat a la qual mesuren (sèries documentals, processos, sistemes d'informació i dades). En el cas de les dades, es desglossen també en els diferents nivells d'agregació conceptual de les dades (dada elemental, objecte, superobjecte i tipus d'objecte). En alguns casos la descripció de l'indicador s'il·lustra amb exemples extrets del cas d'estudi realitzat al Consorci d'Educació de Barcelona.

5.2 Indicadors sobre les sèries documentals

5.2.1 Mètriques descriptives de les sèries documentals²³⁰

Les mètriques descriptives basades únicament en atributs de les sèries documentals són les següents:

SER-01	
Indicador global	Nombre de sèries documentals
Nivells d'agregació	Per caràcter de la sèrie documental
	Per productor de la sèrie documental
	Per tipus de procés generador de la sèrie documental

5.2.2 Mètriques derivades de la relació de la sèrie documental amb altres entitats

Les mètriques descriptives basades en les relacions de les sèries documentals amb altres entitats són les següents:

SER-02	
Indicador de sèrie documental	Nombre d'evidències
Nivells d'agregació	Per tipus d'evidència

L'indicador SER-02 és la base per al càlcul de l'indicador SER-13 Índex de dadificació simple a nivell d'evidència.

En el cas de les diferents regles que regulen el cicle de vida de les sèries documentals, proposem dos indicadors agregats.

SER-03	
Indicador de sèrie documental	Nombre de regles de conservació i eliminació
Nivells d'agregació	Per resolució de la regla

SER-04	
Indicador de sèrie documental	Nombre de regles d'accés i ús
Nivells d'agregació	Per resolució de la regla

²³⁰ L'anàlisi de les mètriques bàsiques de la sèrie documental ja esdevé per sí sol un element de millora de la identificació de sèries en una organització, que pot ser utilitzat en projectes de consultoria en gestió documental, o bé internalitzats. L'anàlisi de les regles de conservació i accés, o de la coherència dels criteris de nominalització, són exemples d'aquesta potencialitat.

Pel que fa a les relacions entre sèries documentals, es proposa el següent indicador descriptiu:

SER-05	
Indicador de sèrie documental	Nombre de sèries documentals relacionades
Nivells d'agregació	Per tipus de relació

Aquest indicador té en compte el nombre de relacions que manté cada sèrie documental amb altres sèries, i el tipus de relació en cada cas. Serveix de base per al càlcul de l'índex d'influència de la sèrie documental sobre altres sèries, que es desenvolupa en els indicadors SER-10, SER-11 i SER-12.

5.2.3 Mètriques sobre les dades de la sèrie documental

En aquest apartat les mètriques es calculen tenint en compte només les ocurrencies de dades en evidències, no les ocurrencies directament imputades a sistemes d'informació que donin suport a sèries documentals.

Les mètriques sobre les dades elementals vinculades a sèries documentals són les que es descriuen amb el següent indicador.

SER-06	
Indicador de sèrie documental	Nombre de dades elementals
Nivells d'agregació	Per tipus de dada
	Per tipus d'evidència
	Per objecte
	Per superobjecte
	Per tipus d'objecte (cosa/acció)

El segon nivell d'anàlisi té en compte el nombre d'ocurrencies de cada dada dins la sèrie documental.

SER-07	
Indicador de sèrie documental	Nombre d'ocurrencies de dades (dades en evidències del procés)
Nivells d'agregació	Per dada elemental
	Per objecte
	Per tipus d'objecte (cosa/acció)
	Per tipus de dada
	Per tipus d'ocurrència (autòctona o importada)
	Per evidència
	Per tipus d'evidència i tipus d'ocurrència (autòctona o importada)

SER-07	
Indicador de sèrie documental	Nombre d'ocurrències de dades (dades en evidències del procés)
	Per tipus d'evidència i objecte
	Per tipus d'evidència i tipus de dada
	Per tipus d'evidència i tipus d'objecte (cosa/acció)

Dues agregacions d'aquest indicador són rellevants per calcular indicadors d'anàlisi. Així, el nombre d'ocurrències per sèrie agregat per tipus d'ocurrència (autòctona o importada) és la base per al càlcul de l'índex de ròssec de la sèrie, recollit en els indicadors SER-16 i SER-17. Per altra banda, el nombre d'ocurrències per sèrie, per tipus d'evidència i tipus d'ocurrència (autòctona o importada) esdevé la base per al càlcul de l'indicador SER-15 (índex de dadificació ponderat).

A partir de les dades es deriven dos indicadors més:

SER-08	
Indicador de sèrie documental	Nombre d'objectes
Nivells d'agregació	Per tipus d'objecte (cosa/acció)

Aquest indicador ens permet determinar tendències pel que fa a la transaccionalitat de la sèrie documental.

SER-09	
Indicador de sèrie documental	Nombre de superobjectes
Nivells d'agregació	Per tipus de superobjecte (cosa/acció)

Aquest indicador és bàsicament una agregació de l'anterior. La utilitat dependrà de com s'hagi desenvolupat la jerarquia d'objectes i superobjectes.

5.2.4 Indicadors d'anàlisi de les sèries documentals

5.2.4.1 Índex d'influència de la sèrie documental

Per avaluar el pes de cada sèrie documental dins el conjunt de l'organització proposem utilitzar dues mètriques combinades.

En primer lloc proposem una mètrica que anomenem **índex d'influència simple**. Aquesta mètrica s'obté de sumar la quantitat de relacions (directes o de primer nivell) que manté cada sèrie amb altres sèries.

SER-10		
Indicador de sèrie documental	Índex d'influència simple	Interpretació de la mesura
Indicadors de base: SER-05 Nombre de sèries documentals relacionades per sèrie documental i tipus de relació	Fórmula de càlcul: [SER-05]	Un nombre elevat de relacions d'una sèrie amb altres sèries indica un potencial d'influència superior d'aquella sèrie. A més relacions, més crítica és la sèrie per al conjunt de l'organització. És per tant un indicador de criticitat .

Però com s'ha detallat en les dues guies metodològiques, les sèries documentals poden mantenir amb altres sèries quatre tipus de relacions. I aquestes relacions no són innòcues: la sèrie antecedent determina fortament la composició i característiques de la sèrie conseqüent, a la qual li transfereix dades com a resultat de l'execució del procés del qual depèn, mentre que una sèrie relacionada no es pot saber en quin nivell la determina. Per al càlcul de l'índex d'influència partim doncs de la inferència que les relacions de continuïtat determinen el grau d'influència d'una sèrie sobre les altres, i els altres tipus de relació no ho fan en la mateixa mesura, o, en tot cas, no es pot determinar. Per incorporar aquesta inferència, hem incorporat la següent ponderació numèrica per cada tipus de relació:

Taula 5-1: Pes en funció del rol de la sèrie documental

Rol de la sèrie	Pes
Sèrie antecedent	-2
Sèrie relacionada	+1
Sèrie conseqüent	+2
Sèrie recapitulativa	+2 ²³¹

Aquest indicador, que anomenem **índex d'influència ponderat**, s'obté d'assignar a cada relació d'una sèrie el pes indicat en la taula anterior, en funció del tipus de relació, i a continuació sumar tots els pesos obtinguts.

²³¹ Per generar un primer indicador, s'ha considerat que les sèries que recapitulen altres sèries tenen el mateix pes que les sèries conseqüents. No obstant això, opinem que cal una reflexió més aprofundida sobre com caldria ponderar les sèries recapitulatives, i si la recapitulació és un tipus de relació o bé un tipus de caràcter de la sèrie. En cas que es consideri un tipus de relació, caldrà escatir si la sèrie recapitulativa realment recapitula (com ha entès tradicionalment l'arxivística), o bé registra per primer cop. Un registre i una taula mestra no funcionen de la mateixa manera: un registre és normalment un tràmit previ a l'inici d'altres activitats, mentre que una taula mestra sempre obtindrà els seus valors de transaccions realitzades anteriorment. Per aquest motiu dediquem un apartat de les conclusions a analitzar si la sèrie recapitulativa pot ser assimilable a la dada mestra d'un model de dades tecnològic (apartat 8.1.5 del capítol 8).

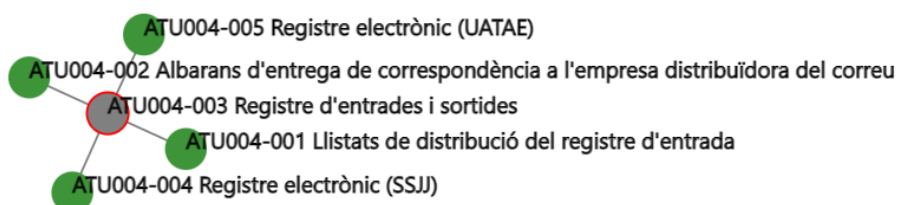
SER-11		
Indicador de sèrie documental	Índex d'influència ponderat	Interpretació de la mesura
Indicadors de base: SER-05 Nombre de sèries documentals relacionades Atribut RELO12-001 Rol de la sèrie	Fórmula de càlcul: [SER-05]*([Pes]Flt[RELO12-001])	Un valor alt en aquest indicador posa de relleu que la sèrie influeix directament sobre la composició i característiques d'altres sèries. És per tant una mesura d' influència .

Aquest indicador, però, pot emascarar la influència real d'una sèrie. Una sèrie amb un nivell d'influència elevat però amb poques relacions amb altres sèries (és a dir amb una criticitat baixa) pot obtenir més puntuació que una sèrie molt crítica, amb influència sobre moltes sèries, però també dependent d'altres sèries.

I a aquest fet cal sumar-hi un segon factor de correcció. Les relacions entre sèries es produeixen en el context d'un mateix procés, no es consignen relacions entre sèries d'altres processos. Però els processos sí que es relacionen entre ells, i adquireixen a través d'aquestes relacions un determinat índex d'influència. Si volem reflectir en l'índex d'influència de la sèrie el seu pes real sobre el conjunt de sèries de l'organització, i no només dins el seu procés, haurem de tenir en compte també l'índex d'influència del procés del qual depèn.

Analitzem un exemple. La sèrie documental ATU004-003 Registre d'entrades i sortides és una sèrie que té un índex d'influència alt dins el seu procés (ATU004 Registre i tramesa de documentació): té quatre relacions amb sèries conseqüents (dos registres interns i dos grups de llistats) que, un cop ponderades i sumades ([SER-10]+[SER-11]) donen un índex d'influència de 12.

Figura 5-3: Mapa de relacions de la sèrie documental ATU004-003



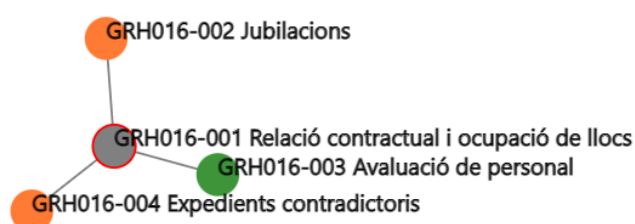
Però pertany a un procés amb un índex d'influència negatiu (-1), atès que és conseqüent d'un altre procés i no té relació amb més processos de l'organització. Es tracta d'un procés de tràmit sense més afectació directa sobre altres processos.

Figura 5-4: Mapa de relacions del procés ATU004



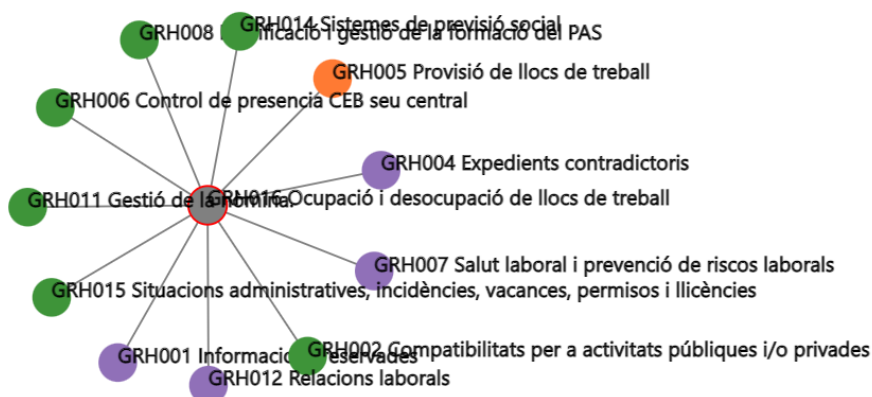
Observem ara un exemple d'una sèrie documental amb un índex d'influència negatiu (-2), producte del fet de tenir dues sèries anteriors i només una sèrie conseqüent. Es tracta de la sèrie GRH016-001 Relació contractual i ocupació de llocs (de treball).

Figura 5-5: Mapa de relacions de la sèrie documental GRH016-001



Però és evident que l'ocupació d'un lloc de treball és el procés que obre la porta a tots els tràmits que posteriorment podrà fer un treballador/a (permisos, llicències, vacances, etc.). Per tant, hem de suposar que és un procés fortament influent sobre la resta de processos de l'organització. Si analitzem les relacions del procés GRH016 Ocupació i desocupació de llocs de treball, es confirma la nostra hipòtesi: l'índex d'influència del procés, un cop ponderades i sumades les relacions ([PRO-11]+[PRO-12]) és de 25.

Figura 5-6: Mapa de relacions del procés GRH016



Per aquest motiu, proposem aplicar un factor de correcció calculant el que anomenem **índex d'influència net**. Aquest indicador el calcularem sumant la influència simple de la sèrie (indicador SER-10) a l'índex d'influència sobre altres sèries (indicador SER-11) i a l'índex d'influència net del procés que genera la sèrie (indicador PRO-13).

SER-12		
Indicador de sèrie documental	Índex d'influència net	Interpretació de la mesura
Indicadors de base: SER-10 Índex d'influència simple SER-11 Índex d'influència ponderat PRO-13 Índex d'influència net	Fórmula de càlcul: [SER-10]+[SER-11]+[PRO-13]	Un valor elevat en aquest índex indica que la sèrie documental influeix a través de les seves relacions a altres sèries documentals del conjunt de l'organització, i en pot determinar aspectes de contingut i característiques.

Si apliquem aquest factor de correcció al cas anterior, obtenim el següent resultat:

Taula 5-2: Càlcul de l'índex d'influència sense i amb factor de correcció

	Sense factor de correcció			Amb factor de correcció			
	SER-10	SER-11	SER-12	SER-10	SER-11	PRO-13	SER-12
ATU004-003 Registre d'entrades i sortides	4	8	12	4	8	-1	11
GRH016-001 Relació contractual i ocupació de llocs	3	-2	1	3	-2	25	26

5.2.4.2 Índex de dadificació de la sèrie documental

Aquest és un indicador especialment rellevant per avaluar l'avenç del procés de transformació digital en una organització. El calculem en diferents versions.

L'índex de dadificació simple a nivell d'evidència correspon al nombre d'evidències de la sèrie documental que són pantalles d'entrada i sortida de dades, i que estan categoritzades com de tipus "Dades" en l'atribut ENT004-005.

SER-13		
Indicador de sèrie documental	Índex de dadificació simple a nivell d'evidència	Interpretació de la mesura
Indicadors de base: SER-02 Nombre d'evidències per tipus d'evidència ENT004-003 Tipus d'evidència	Fórmula de càlcul: ([SER-02]Fit[ENT004-003]="Dades") / [SER-02]	Un percentatge elevat indica una sèrie amb una part important d'accions que s'emmagatzemen en sistemes d'informació i generen dades estructurades.

Una forma possiblement més precisa consisteix en calcular l'índex de dadificació simple a nivell de dada elemental. En aquest cas es calcula quina part del total de dades elementals de la sèrie apareixen en evidències de tipus dada, i per tant estan recollides en sistemes d'informació.

Una sèrie documental amb un nombre elevat de documents-objecte tindrà un índex de dadificació a nivell d'evidència molt baix, amb independència de com s'hagin generat aquestes evidències. Si les evidències-objecte recullen dades que no s'acaben emmagatzemant en cap sistema d'informació, com s'esdevé en els processos on les evidències es generen manualment, el nivell de dadificació serà realment baix. Però si la major part de les dades que les evidències-objecte recullen també la trobem en els sistemes d'informació que formen part de la sèrie documental, com s'esdevé en els processos on els documents-objecte es generen a partir d'aplicacions informàtiques i plantilles predefinides, aleshores el nivell de dadificació serà alt. Aquest indicador actua doncs de corrector i alternativa de l'anterior (SER-13).

SER-14		
Indicador de sèrie documental	Índex de dadificació simple a nivell de dada elemental	Interpretació de la mesura
Indicadors de base: SER-06 Nombre de dades elementals ENT004-003 Tipus d'evidència	Fórmula de càlcul: ([SER-06]Fit[ENT004-003]="Dades") / [SER-06]	Un percentatge elevat indica una sèrie amb una part important d'accions que s'emmagatzemen en sistemes d'informació i generen dades estructurades.

Però l'índex de dadificació basat en dades només expressa la potència que pot tenir una sèrie documental de gestionar una part de la seva informació en forma de dades estructurades, no demostra pas que ho estigui fent. És a dir, en una sèrie documental pot existir una reutilització intensiva per part de diferents documents-objecte d'una dada que existeix en un sistema d'informació, sense que existeixi cap mena de sincronització entre les ocurrències d'aquesta dada en cadascuna de les evidències. La dada es pot estar introduint manualment a cada evidència, cosa la qual indica que no s'estaran explotant els beneficis de la dadificació. Per a que aquesta dadificació operi de forma activa, cal que els sistemes d'informació de la sèrie documental esdevinguin els proveïdors de dades a la resta d'evidències, i especialment a les evidències-objecte.

Per poder reflectir aquest nivell d'integració en l'índex de dadificació incorporarem un element: la forma d'aparició de la dada en cada evidència. Com s'ha explicat a la **GDAD** (apartat 3.3.6.2.3 del capítol 3), si la dada apareix per primer cop en aquella evidència, o requereix que s'informi el seu valor per completar l'evidència, la considerarem una dada "autòctona" de l'evidència, mentre que si la dada apareix amb un valor ja informat la considerarem una dada "importada".

Utilitzant aquestes dues categories calcularem l'índex de dadificació ponderat. En aquest cas, s'assignarà a cada dada elemental un pes en funció de les característiques de l'evidència on apareix i del tipus d'ocurrència. L'objectiu és donar més valor a les ocurrencies importades, que indiquen que existeix un nivell major d'integració o connectivitat entre les evidències. I en segon lloc, donar també valor a les evidències que estan en forma de sistemes d'informació. La ponderació que es proposa és la que s'indica a la taula següent:

Taula 5-3: Pes en funció del tipus i lloc de la ocurrència

Ocurrència de dada elemental	Pes
Apareix en una evidència tipus "Dades" com importada	2
Apareix en una evidència tipus "Dades" com autòctona	1
Apareix en una evidència tipus "Document" com importada	1
Apareix en una evidència tipus "Document" com autòctona	0

Per evitar un indicador amb valors superiors a l'1, aquest valor ponderat de cada dada elemental es contrastarà amb el valor teòric en un entorn totalment dadificat, en el qual es planteja que cada dada apareix un sol cop com autòctona en una evidència de tipus dades, i després com importada en la resta d'evidències, també de tipus dades. Per tant, el valor de contrast de la suma de les ocurrencies reals ponderades serà el nombre d'ocurrencies en la sèrie documental ponderat d'acord amb la taula de valors anterior: (total ocurrencies*2)-1. La divisió del valor real entre el valor teòric calculat segons aquest patró serà el valor de l'índex de dadificació ponderat.

SER-15		
Indicador de sèrie documental	Índex de dadificació ponderat	Interpretació de la mesura
Indicadors de base: SER-07 Nombre d'ocurrencies de dades per tipus d'evidència i tipus d'ocurrència ENT004-003 Tipus d'evidència RELO20-001 Tipus d'ocurrència	Fórmula de càlcul: $\frac{([SER-07]*([Pes]Flt([ENT004-003] AND [RELO20-001])))}{([SER-07]*2)-1}$	Un valor alt en l'índex de dadificació ponderat indica un ús eficient de les dades dins de la sèrie documental i un pes important dels sistemes d'informació com a base de la tramitació.

5.2.4.3 Índex de ròssec²³²

²³² Per donar nom a aquest indicador hem utilitzat la terminologia pròpia de la metodologia APS d'anàlisi i millora de circuits documentals (Puig-Pey & Guiu, 2002, 2008, 2013), on es diferencia entre variables autòctones i de ròssec. En la **GDAD** (apartat 3.3 del capítol 3) hem utilitzat el terme "importades" per a definir aquestes variables, per la qual cosa podríem utilitzar també per

L'índex de ròssec és una mesura de reutilització de les dades en una sèrie documental. Parteix de la hipòtesi que, en un circuit evidencial òptim, cada dada s'informa només un sol cop, i la resta de cops és reutilitzada amb el mateix valor. A diferència de l'índex de dadificació ponderat (indicador SER-15), en aquest indicador no es té en compte el tipus ni format de l'evidència on apareix cada dada. Tampoc s'inclou en el supòsit teòric l'existència de variables que modifiquen el valor en el curs de l'execució del procés (com s'esdevé per exemple amb l'estat de l'expedient).

La importació de dades (o ròssec) es pot manifestar a dos nivells:

- Ròssec de variables entre les evidències d'una mateixa sèrie documental, el qual anomenarem **índex de ròssec intrasèrie**.
- Ròssec de variables entre sèries documentals relacionades, que anomenarem **índex de ròssec intersèrie**.

5.2.4.3.1 Ròssec intrasèrie

L'índex de ròssec intrasèrie parteix, com hem comentat, del supòsit teòric que tota dada s'informa en una sèrie documental en una primera evidència (com autòctona d'aquella evidència), i a partir d'aquí exporta el seu valor a la resta d'evidències de la sèrie que n'hagin de menester (com importada en la resta d'evidències). Així doncs, l'índex de ròssec teòric de les dades d'una sèrie és igual al nombre d'ocurrències de cada dada menys 1.

Per obtenir l'índex, haurem de dividir el nombre d'ocurrències de cada dada que tenen la condició d'importades pel nombre teòric d'ocurrències importades que correspondrien a aquella dada. Per evitar distorsions en l'indicador descartarem aquelles dades que només tenen una ocurrència, restant 1 al valor divisor.

SER-16		
Indicador de sèrie documental	Índex de ròssec intrasèrie	Interpretació de la mesura
Indicadors de base: SER-07 Nombre d'ocurrències de dades per dada elemental i tipus d'ocurrència RELO20-001 Tipus d'ocurrència	Fórmula de càlcul: $([\text{SER-07}]_{\text{Fit}}[\text{RELO20-001}] = \text{"IMP"}) / ([\text{SER-07}]-1)$	Un valor proper a 1 d'aquest índex indica una alta reutilització de cada dada i un desenvolupament òptim del procés o tramitació.

aquest indicador el nom "índex d'importació", i també volem proposar com alternatives els noms de "índex de dependència dadifical", o bé "índex d'herència de dades". En tot cas, mantenim en aquesta versió el nom "índex de ròssec".

Aquest indicador es calcula a nivell de dada elemental, partint de la matriu de variables del procés i de la sèrie documental. Agafem l'exemple de la sèrie INF001-003 Elaboració d'estudis amb col·laboracions externes, la qual forma part del procés INF001 Gestió, manteniment i ús del Banc de Dades del CEB. Aquesta sèrie es compon de quatre evidències:

Taula 5-4: Evidències de la sèrie documental INF001-003

EVIDÈNCIA	FORMAT
INF001-E012 Petició d'informe, estudi i projecte de recerca en dades	Correu electrònic
INF001-E013 Indicació de les condicions del projecte de recerca	Correu electrònic
INF001-E014 Registre de l'informe, estudi o projecte de recerca	Biblioteca de Sharepoint
INF001-E015 Informe o estudi de recerca	PDF

I la matriu de variables d'aquesta sèrie es disposa de la forma següent:

Taula 5-5: matriu de variables de la sèrie documental INF001-003

DADES	INF001-E012	INF001-E013	INF001-E014	INF001-E015
Nom unitat administrativa	AUT	AUT	AUT	
Observacions - informe	AUT	AUT	AUT	
Nom i cognoms treballador CEB	AUT	IMP	IMP	
Correu electrònic treballador CEB	AUT	IMP		
Servei o projecte	AUT		AUT	
Nom i cognoms persona externa		AUT		IMP
Correu corporatiu		AUT		
Data de l'informe			AUT	AUT
Nom de l'informe			AUT	AUT
Revisat			AUT	IMP
Abast geogràfic			AUT	
Descriptors estudis			AUT	
Etapa educativa			AUT	
Publicat Web			AUT	
Correu electrònic entitat				AUT
NIF entitat				AUT
Nom entitat col·laboradora				AUT

El càlcul de l'índex de ròssec de la sèrie documental INF001-003 Elaboració d'estudis amb col·laboracions externes dona com a resultat un índex de 0,18. Aquest índex, desglossat a nivell de dada, només té valor en quatre dades:

Taula 5-6: Ròssec individualitzat per dada elemental de la sèrie documental INF001-003

DADES	RÒSSEC
Nom i cognoms treballador CEB	2
Correu electrònic treballador CEB	1
Nom i cognoms persona externa	1
Revisat	1

Un element a tenir present, i que condiciona la validesa de l'índex de ròssec teòric, és la pauta de marcatge que estableixi la metodologia de recollida de dades per a la primera ocurrència d'una dada en una sèrie. Pot haver-hi casos, com comentem en la descripció del ròssec intersèrie, on la primera ocurrència de la dada contingui un valor que s'importa d'un procés o sèrie antecedent, amb la qual cosa es pot donar el cas que una dada no tingui cap ocurrència autòctona en una sèrie documental. En aquest cas, si la dada està classificada com "importada" en totes les ocurrències, el model de ròssec teòric abans exposat no s'hauria d'aplicar, i el ròssec teòric hauria de ser únicament el nombre total d'ocurrències. En el cas del grup focal, la pauta metodològica ha estat marcar com a autòctona la primera ocurrència de tota dada, amb la qual es manté la validesa de l'indicador SER-16 tal com s'ha definit anteriorment.

5.2.4.3.2 Ròssec intersèrie

Un dels objectius de la dadificació dels processos és aconseguir, a partir de l'estructuració de les dades d'un procés, que aquestes dades es reutilitzin en d'altres processos, i es redueixi la duplicació d'informació. El ròssec o reutilització de variables entre sèries documentals, i en conseqüència entre processos, és un indicador d'integració de la informació a nivell de tota l'organització. Un índex elevat entre sèries indica un elevat nivell de reaprofitament de la informació en una organització, i pot ser indicador d'una major proximitat a l'ideal de la dada única.

Aquest indicador es pot calcular de dues maneres. La primera, més simple, no té en compte el tipus d'ocurrència de la dada (autòctona o importada), sinó només la coincidències de la mateixa dada entre dues sèries documentals. És per tant un indicador d'integració en potència, però no permet determinar si aquesta integració es materialitza. Per calcular aquest indicador, cal realitzar tres operacions:

1. A partir de la matriu de variables, obtenir el diccionari de cada sèrie documental.
2. A continuació recuperar les relacions entre sèries documentals recollides al mapa.
3. I en tercer lloc comparar els diccionaris de variables de les sèries relacionades, i obtenir, a partir de les dades coincidents sobre el total de dades elementals d'ambdues sèries, l'índex de ròssec intersèrie.

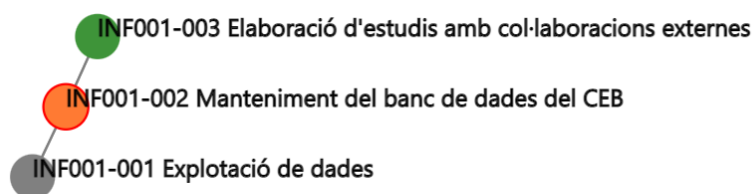
Aquest indicador, per tant, es calcula sempre sobre la intersecció de dues sèries.

SER-17		
Indicador de sèrie documental	Índex de ròssec intersèrie	Interpretació de la mesura
Indicadors de base: SER-06 Nombre de dades elementals	Fórmula de càlcul: [SER-06] de [Sèrie 1] AND [Sèrie 2] / [SER-06] de [Sèrie 1] OR [Sèrie 2]	Un valor elevat d'aquest indicador indica la possibilitat de reutilitzar dades entre dues sèries, i

SER-17		
Indicador de sèrie documental	Índex de ròssec intersèrie	Interpretació de la mesura
		en conseqüència de compartir o connectar sistemes d'informació.

Per analitzar un exemple, ens centrarem en tres sèries documentals relacionades entre elles:

Figura 5-7: Relacions entre tres sèries documentals del procés INF001



Les relacions que mantenen aquestes sèries entre elles són les següents:

Taula 5-7: Tipus de relacions entre tres sèries documentals del procés INF001

SÈRIE DOCUMENTAL	RELACIÓ	SÈRIE DOCUMENTAL
INF001-002 Manteniment del banc de dades del CEB	És antecedent de	INF001-001 Explotació de dades
INF001-002 Manteniment del banc de dades del CEB	És antecedent de	INF001-003 Elaboració d'estudis amb col·laboracions externes

L'índex de ròssec d'aquestes dues relacions és el següent:

Taula 5-8: Índex de ròssec intersèrie de les relacions entre tres sèries documentals del procés INF001

SÈRIE DOCUMENTAL	SÈRIE DOCUMENTAL	ÍNDEX DE RÒSSEC
INF001-002 Manteniment del banc de dades del CEB	INF001-001 Explotació de dades	0,06
INF001-002 Manteniment del banc de dades del CEB	INF001-003 Elaboració d'estudis amb col·laboracions externes	0,07

Observem que, tot i tractar-se de sèries documentals relacionades directament, l'índex de ròssec intersèrie és molt baix. Més tenint present que dues de les sèries comparteixen un sistema d'informació: el sistema SI02.012 Catàleg de processos del CEB.

Taula 5-9: Relacions entre tres sèries documentals i els sistemes d'informació del procés INF001

SÈRIE DOCUMENTAL	SISTEMA D'INFORMACIÓ
INF001-001 Explotació de dades	SI08.021 INF001-001 Peticions d'informació del Banc de Dades
INF001-002 Manteniment del banc de dades del CEB	SI02.012 Catàleg de processos del CEB
	SI08.022 INF001-002 Manteniment del banc de dades
INF001-003 Elaboració d'estudis amb col·laboracions externes	SI08.023 SINF001-007 Funcionament del Gabinet d'estudis del CEB

Però si ampliem l'abast a la resta de sèries del procés INF001, observem un nombre major de coincidències:

Taula 5-10: Índex de ròssec intersèrie de les relacions entre totes les sèries documentals del procés INF001

SÈRIE DOCUMENTAL	SÈRIE DOCUMENTAL	ÍNDEX DE RÒSSEC
INF001-001 Explotació de dades	INF001-002 Manteniment del banc de dades del CEB	0,07
INF001-001 Explotació de dades	INF001-003 Elaboració d'estudis amb col·laboracions externes	0,38
INF001-001 Explotació de dades	INF009-001 Manteniment de les dades de la Base de Dades de Centres (BDC)	0,18
INF001-001 Explotació de dades	INF009-002 Manteniment de les dades de la Base de dades de Centres (BDC)	0,29
INF001-002 Manteniment del banc de dades del CEB	INF001-001 Explotació de dades	0,07
INF001-002 Manteniment del banc de dades del CEB	INF001-003 Elaboració d'estudis amb col·laboracions externes	0,07
INF001-002 Manteniment del banc de dades del CEB	INF009-001 Manteniment de les dades de la Base de Dades de Centres (BDC)	0,21
INF001-002 Manteniment del banc de dades del CEB	INF009-002 Manteniment de les dades de la Base de dades de Centres (BDC)	0,18
INF001-003 Elaboració d'estudis amb col·laboracions externes	INF001-001 Explotació de dades	0,38

SÈRIE DOCUMENTAL	SÈRIE DOCUMENTAL	ÍNDIX DE RÒSSEC
INF001-003 Elaboració d'estudis amb col·laboracions externes	INF001-002 Manteniment del banc de dades del CEB	0,07
INF001-003 Elaboració d'estudis amb col·laboracions externes	INF009-001 Manteniment de les dades de la Base de Dades de Centres (BDC)	0,20
INF001-003 Elaboració d'estudis amb col·laboracions externes	INF009-002 Manteniment de les dades de la Base de dades de Centres (BDC)	0,33

Podem observar doncs una possible relació no prevista inicialment entre dues sèries del mateix procés (INF001-001 i INF001-003). I en el segon cas, observem com dues sèries d'aquest procés tenen una relació amb una altra sèrie (INF009-002), que emmagatzema les seves dades en un sistema d'informació consumit habitualment pels dos processos que s'hi relacionen, el sistema SI02.001 Base de Dades de Centres (BDC).

Cal tenir present que quan les dades d'una sèrie documental són molt genèriques (dades identificatives i de contacte d'alumne, de treballador/a, etc.) es poden produir falses coincidències entre sèries que no mantenen cap relació funcional entre elles (de processos i famílies diferents). Aquest fenomen respon el que es denomina "Relació de contingut" entre sèries, i afecta especialment als objectes de tipus "Cosa".²³³ Per aquest motiu, i amb l'objectiu de reduir les falses coincidències quan no existeix relació funcional entre les sèries documentals, es planteja l'opció de limitar les coincidències als objectes transaccionals (de tipus "Acció").

Aquest indicador es pot calcular també en una segona modalitat, que consisteix en incorporar el tipus d'ocurrència en la fórmula de càlcul. Utilitzant ponderacions es pot atribuir un pes major a una relació en la que el procés conseqüent té una primera aparició d'una variable en la modalitat "Importada", el que pot indicar una relació de transferència de dades entre sistemes d'ambdues sèries documentals, o la compartició funcional d'un mateix sistema. Aquesta segona modalitat requereix, però, marcar variables inicials de sèrie com a importades en la primer ocurrència, opció que no hem aplicat en el nostre projecte. Per aquest motiu, no incloem exemples d'aquest indicador.

L'índex de ròssec intersèrie, més enllà de la descripció i mesura de la intensitat de les relacions entre sèries documentals vinculades funcionalment, té un potencial important com a recurs per a l'avaluació i millora de les relacions entre entitats del mapa de sèries documentals i de sistemes d'informació. L'anàlisi de l'índex de ròssec pot facilitar:

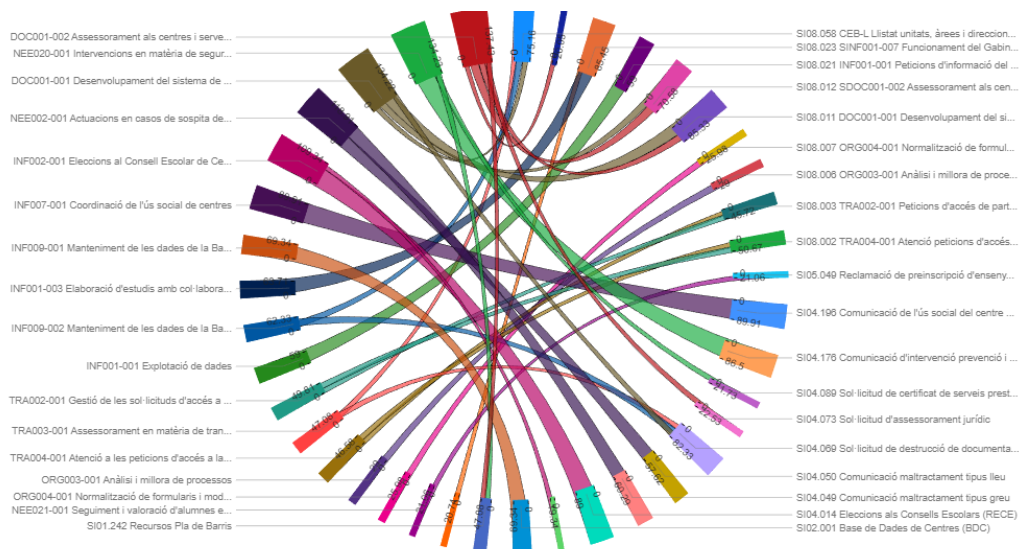
- La detecció de relacions funcionals (de continuïtat o dependència) entre sèries documentals que els professionals de l'organització no hagin pogut detectar

²³³ Vegeu la definició a la **GMAP**, a la **GDAD** i al model de dades (capítols 3 i 4).

durant el treball de camp, especialment quan la coincidència afecti dades d'objectes transaccionals (del tipus "Acció").

- La detecció d'oportunitats per reutilitzar sistemes d'informació com a suport de diferents sèries documentals, o bé per connectar sistemes que comparteixen les mateixes dades, facilitant en aquest cas un primer embrió del model de dades d'aquest intercanvi.
- La detecció de sèries documentals de caràcter recapitulatiu, quan una sèrie documental concentri dades de nombroses sèries d'origens diferents. I, de la mateixa manera, a identificar possibles sistemes d'informació candidats a sistemes gestors de dades mestres.²³⁴
- La detecció de relacions de contingut entre sèries documentals que permetin optimitzar i homogeneïtzar els criteris de nominalització de les sèries que comparteixen objectes descrits, element recollit en els objectius instrumentals.

Figura 5-8: Exemple d'anàlisi visual de coincidències en dades elementals entre sèries documentals i sistemes d'informació



5.2.4.4 Índex de transaccionalitat de la sèrie documental

L'índex de transaccionalitat de la sèrie documental té per objectiu recollir el volum d'informació de la sèrie documental que té caràcter o significació transaccional, és a dir que està conceptualment vinculat a accions o fets. Aquest indicador es proposa en tres variants:

²³⁴ En cas que es confirmi la hipòtesi de l'objectiu 5 (apartat 8.1.5 del capítol 8).

5.2.4.4.1 Índex de transaccionalitat calculat a partir de l'objecte de les dades

Amb aquest indicador es pretén representar el caràcter majoritari de les dades de la sèrie documental. Aquest indicador es calcula a partir de les dades elementals presents en les evidències de cada sèrie, sense tenir en compte el nombre d'ocurrències de cada dada. Consisteix en la proporció de dades elementals de la sèrie documental l'objecte de les quals és equiparable a acció sobre la matriu global de dades de la sèrie documental.

SER-18		
Indicador de sèrie documental	Índex de transaccionalitat a partir de les dades elementals	Interpretació de la mesura
Indicadors de base: SER-06 Nombre de dades elementals per sèrie documental i tipus d'objecte (cosa/acció) ENT014-002 Tipus d'objecte	Fórmula de càlcul: ([SER-06]Flt[ENT014-002]="Acció") / [SER-06]	Un nombre elevat de dades descriptives d'objectes de tipus "acció" indicarà un nivell més alt de transaccionalitat de la sèrie.

5.2.4.4.2 Índex de transaccionalitat calculat a partir del tipus de dades

També podem calcular aquest índex tenint en compte únicament el tipus de dada. Aquesta segona modalitat de comptatge, malgrat ser més simple, és menys precisa, atès que no té en compte aquelles dades de tipus no transaccional que descriuen objectes de tipus "acció". La fórmula de càlcul és la següent:

SER-19		
Indicador de sèrie documental	Índex de transaccionalitat a partir del tipus de dades	Interpretació de la mesura
Indicadors de base: SER-06 Nombre de dades elementals per sèrie documental ENT006-003 Tipus de dada	Fórmula de càlcul: ([SER-06]Flt[ENT006-003]="Transaccional") / [SER-06]	Un valor elevat d'aquest índex indica un volum important de gestió sobre el conjunt de la matriu de dades de la sèrie documental.

Com es detalla en l'explicació de l'índex SIS-18, l'equivalent a nivell de sistema d'informació, aquest índex és propens a donar falsos negatius, atès que el pes de les dades transaccionals pot aparèixer sobrerrepresentat en una sèrie on es recullin poques dades de les entitats tipus "cosa", sense que això impliqui una major transaccionalitat real.

5.2.4.4.3 Índex de transaccionalitat calculat a partir dels objectes

Aquest indicador és una variant de l'indicador SER-18, centrat només en el comptatge dels tipus d'objectes descrits en la matriu de dades de la sèrie documental. Es calcula com segueix:

SER-20		
Indicador de sèrie documental	Índex de transaccionalitat a partir dels objectes	Interpretació de la mesura
Indicadors de base: SER-08 Nombre d'objectes per sèrie documental ENT014-002 Tipus d'objecte	Fórmula de càlcul: ([SER-08]Fit[ENT014-002]="Acció") / [SER-08]	Un índex elevat indica una vinculació major de la sèrie documental amb accions, i per tant una major probabilitat d'esdevenir sèrie transaccional.

Cal tenir present que el nombre d'objectes de tipus "cosa" que intervinguin en l'acció podrà incidir i, en tot cas, distorsionar aquest índex. Una activitat que impliqui la participació de molts agents diferents obtindrà un índex més baix que una altra activitat on el nombre d'agents sigui baix, malgrat ser ambdues activitats vinculades amb una acció i amb un caràcter transaccional obvi.

5.2.5 Mapa de relacions entre sèries documentals

Per a l'anàlisi visual de les relacions entre les sèries documentals es poden generar visualitzacions que mostrin de forma interactiva totes les relacions entre sèries documentals. Per un exemple vegeu el cas d'aplicació al Consorci d'Educació de Barcelona a l'apartat 6.4.5 del capítol 6.

5.3 Indicadors sobre els processos

5.3.1 Mètriques descriptives dels processos

Les mètriques descriptives basades únicament en atributs dels processos són les següents:

PRO-01	
Indicador global	Nombre de processos
Nivells d'agregació	Per tipus de procés
	Per família de processos

5.3.2 Mètriques derivades de les relacions del procés amb altres entitats

Les mètriques descriptives basades en les relacions dels processos amb altres entitats són les següents:

PRO-02	
Indicador de procés	Nombre de tasques
Nivells d'agregació	Per tipus de tasca (manual o automatitzada)
	Pel rol de la tasca en el flux

Aquesta mètrica, agregada en funció dels tipus de tasca, obre la porta a dos possibles indicadors. En el primer cas, la diferenciació en funció del tipus de tasca permet disposar d'una mesura del nivell d'automatització dels processos i les famílies.

En el segon cas, la tipologia específica o rol de la tasca dins els processos també pot donar una idea de la complexitat del procés. Per exemple, un procés amb un nombre elevat de bifurcacions serà més complex en el seu circuit documental que un procés seqüencial, mentre un procés amb un nombre elevat d'esdeveniments tindrà una dependència més gran d'entrades externes.

PRO-03	
Indicador de procés	Nombre d'evidències
Nivells d'agregació	Per tipus d'evidència

Quan es considera el tipus genèric d'evidència (document-objecte o dades estructurades), aquest indicador esdevé la base per al càlcul de l'índex de dadificació del procés (indicador PRO-14 a 16).

PRO-04	
Indicador de procés	Nombre de sèries documentals
Nivells d'agregació	Per caràcter de la sèrie documental

El nombre de sèries documentals per procés és una mesura fortament influenciada pel criteri aplicat pels professionals en el procés d'identificació de les sèries, i que no es pot desvincular del tipus de sèrie (primària o secundària) i del rol que adopten en les relacions entre elles.

Una altra relació que té potencial significatiu són els sistemes d'informació que donen suport als processos.

PRO-05	
Indicador de procés	Nombre de sistemes d'informació
Nivells d'agregació	Per entorn del sistema d'informació
	Per tecnologia

l'indicador que reflecteix el rol d'un procés dins el mapa de processos és el nombre de processos amb els quals es relaciona, i de quina forma s'hi relaciona.

PRO-06	
Indicador de procés	Nombre de processos relacionats
Nivells d'agregació	Per tipus de relació

El nombre de processos relacionats és un indicador bàsic per al càlcul de l'índex d'influència del procés, que es detalla en els indicadors PRO-11, PRO-12 i PRO-13.

5.3.3 Mètriques sobre les dades del procés

En aquest apartat les mètriques es calculen tenint en compte només les ocurrències en evidències, no les ocurrències directament imputades a sistema d'informació.

El primer indicador té en compte el nombre de dades elementals que apareixen en evidències vinculades al procés.

PRO-07	
Indicador de procés	Nombre de dades elementals
Nivells d'agregació	Per objecte
	Per superobjecte
	Per tipus d'objecte (cosa/acció)
	Per tipus de dada
	Per tipus d'evidència

Aquest indicador és la base per al càlcul dels índex de dadificació a nivell de dada elemental recollits en l'indicador PRO-15.

El segon indicador té en compte el nombre d'ocurrències de cada dada elemental en evidències vinculades al procés.

PRO-08	
Indicador de procés	Nombre d'ocurrències de dades (dades en evidències del procés)
Nivells d'agregació	Per dada elemental
	Per objecte
	Per superobjecte
	Per tipus d'objecte (cosa/acció)
	Per tipus de dada
	Per tipus d'ocurrència (autòctona o importada)
	Per evidència
	Per tipus d'evidència i tipus d'ocurrència (autòctona o importada)

Aquest indicador, agregat per tipus d'ocurrència, és la base per al càlcul de l'índex de ròssec del procés, recollit en els indicadors PRO-17 i PRO-18. Agregat per tipus d'evidència i tipus d'ocurrència també ens servirà per calcular l'indicador PRO-16.

Finalment, es proposen dos indicadors complementaris, a nivell d'objecte i superobjecte, que permeten valorar el pes de cada objecte dins el procés. Els patrons d'objectes per procés s'analitzen a les conclusions.

PRO-09	
Indicador de procés	Nombre d'objectes
Nivells d'agregació	Per tipus d'objecte (cosa/acció)

PRO-10	
Indicador de procés	Nombre de superobjectes
Nivells d'agregació	Per tipus de superobjecte (cosa/acció)

5.3.4 Indicadors d'anàlisi dels processos

5.3.4.1 Índex d'influència del procés

Per avaluar el pes de cada procés dins el conjunt de l'organització, utilitzem en primer lloc una mètrica que anomenem índex d'influència simple. Aquest índex representa la intensitat amb la qual un procés influeix sobre altres processos de l'organització.

Aquesta mètrica s'obté de sumar la quantitat de relacions (directes o de primer nivell) que manté cada procés amb altres processos.

PRO-11		
Indicador de procés	Índex d'influència simple	Interpretació de la mesura
Indicadors de base: PRO-06 Nombre de processos relacionats	Fórmula de càlcul: [PRO-06]	Un nombre elevat de relacions amb altres processos indica un potencial d'influència superior sobre l'activitat de l'organització. És per tant un indicador de criticitat .

Però com s'ha detallat en la **GMAP**, els processos poden mantenir amb altres processos tres tipus de relacions (apartat 3.2.4.4 del capítol 3). Per al càlcul de l'índex d'influència partim de la inferència que el procés antecedent determina fortament al procés conseqüent, mentre que un procés relacionat o depenent només el determina mitjanament. Per incorporar aquesta inferència, hem incorporat la següent ponderació numèrica per a determinar el pes de cada tipus de relació:

Taula 5-11: Pes en funció del rol del procés

Rol del procés	Pes influència
Procés antecedent	2
Procés depenent	0
Procés conseqüent	-2

Aquest indicador, que anomenem índex d'influència ponderat, s'obté d'assignar a cada relació d'un procés el pes indicat en la taula anterior, en funció del tipus de relació, i a continuació sumar tots els pesos obtinguts.

PRO-12		
Indicador de procés	Índex d'influència ponderat	Interpretació de la mesura
Indicadors de base: PRO-06 Nombre de processos relacionats per tipus de relació RELO08-001 Relació de continuïtat	Fórmula de càlcul: [PRO-06]*([Pes]Flt[RELO08-001])	Un valor alt en aquest indicador posa de relleu que el procés determina l'execució d'altres processos. És una mesura d' influència .

Aquest indicador, però, pot emmascarar la importància real d'un procés. Per aquest motiu, per calcular el que anomenem índex d'influència net, l'operació que farem és sumar l'índex d'influència ponderat a l'índex d'influència simple. Aquest serà l'indicador que agafarem com a referència.

PRO-13		
Indicador de procés	Índex d'influència net	Interpretació de la mesura
Indicadors de base: PRO-11 Índex d'influència simple PRO-12 Índex d'influència ponderat	Fórmula de càlcul: [PRO-11]+[PRO-12]	Com en l'indicador PRO-12, un valor alt en aquest indicador posa de relleu que el procés determina l'execució d'altres processos, o bé en depèn.

Aquest indicador es pot millorar si considerem no només les relacions directes o de primer nivell, sinó la cadena completa de relacions, però només en el cas que siguin determinants (relacions de conseqüència o dependència).

Els processos amb un índex d'influència més alt són els processos determinants de l'organització, aquells que condicionen l'execució de la resta de processos. Els processos amb un índex d'influència més baix són, en canvi, processos menys crítics per a l'organització, i en tot cas relacionats amb activitats d'avaluació o de suport no imprescindible.

També es pot convenir que els processos amb un índex d'influència alt són els processos origen, aquells que obren portes funcionals a la resta de processos. Poden ser processos amb poc pes d'execució, com és la planificació estratègica, però per l'impacte que tenen en altres processos es poden arribar a considerar crítics.

5.3.4.2 Índex de dadificació del procés

Aquest és un indicador especialment rellevant per avaluar l'avenç del procés de transformació digital en una organització. Es calcula en diferents versions:

L'índex de dadificació simple a nivell d'evidència del procés correspon al percentatge d'evidències del procés que són pantalles d'entrada i sortida de dades, és a dir categoritzades com de tipus "dades" en l'atribut ENT004-005.

PRO-14		
Indicador de procés	Índex de dadificació simple a nivell d'evidència	Interpretació de la mesura
Indicadors de base: PRO-03 Nombre d'evidències per tipus d'evidència ENT004-003 Tipus d'evidència	Fórmula de càlcul: $([PRO-03]Flt[ENT004-003]="Dades") / [PRO-03]$	Un percentatge elevat indica un procés amb una part important d'accions que es realitzen sobre sistemes d'informació i generen dades estructurades.

També es pot calcular l'**índex de dadificació simple a nivell de dada elemental**. En aquest cas es calcula el percentatge sobre el total de dades elementals del procés de dades que apareixen en dos tipus d'entitats: evidències tipus dada, i sistemes d'informació.

En el primer cas, es calcula el percentatge, sobre totes les dades elementals del procés, de dades que apareixen en algun moment en evidències de tipus dada.

PRO-15		
Indicador de procés	Índex de dadificació simple a nivell de dada elemental	Interpretació de la mesura
Indicadors de base: PRO-07 Nombre de dades elementals ENT004-003 Tipus d'evidència	Fórmula de càlcul: $([PRO-07]Flt[ENT004-003]="Dades") / [PRO-07]$	Un percentatge elevat indica, com en l'indicador anterior, que una part important de les accions del procés es realitzen sobre sistemes d'informació i generen dades estructurades.

Finalment, amb l'**índex de dadificació ponderat** s'incorpora un factor d'afinament en els percentatges anteriors, especialment en aquells casos en què una mateixa dada és reutilitzada amb freqüència dins el procés i apareix simultàniament en evidències de tipus dada i document. En aquest cas, s'assigna a cada dada elemental un pes en funció de les característiques de l'evidència on apareix i del tipus d'ocurrència, tal com s'indica en la taula següent:

Taula 5-12: Pes en funció del tipus i lloc de la ocurrència

Ocurrència de dada elemental	Pes assignat
Apareix en una evidència tipus "Dada" com autòctona	2
Apareix en una evidència tipus "Dada" com importada	1
Apareix en una evidència tipus "Document" com importada	1
Apareix en una evidència tipus "Document" com autòctona	0

Com hem fet amb les sèries documentals, per evitar un indicador amb valors superiors a l'1, aquest valor ponderat de cada dada elemental es contrastarà amb el valor teòric en un entorn totalment dadificat, en el qual es planteja que cada dada apareix un sol cop com autòctona en una evidència de tipus dada, i després com importada en la resta d'evidències. Per tant, el valor de contrast de la suma de les ocurrències reals ponderades serà el nombre d'ocurrències en el procés, però ponderat d'acord amb la taula de valors anterior: (total ocurrències*2)-1. La divisió del valor real entre el valor teòric calculat segons aquest patró serà el valor de l'índex de dadificació ponderat.

PRO-16		
Indicador de procés	Índex de dadificació ponderat	Interpretació de la mesura
Indicadors de base: PRO-08 Nombre d'ocurrències de dades per tipus d'evidència i tipus d'ocurrència ENT004-003 Tipus d'evidència REL020-001 Tipus d'ocurrència	Fórmula de càlcul: $\frac{([PRO-08]*([Pes]Flt([ENT004-003] AND [REL020-001])))}{(([PRO-08]*2)-1)}$	Un valor alt en l'índex de dadificació ponderat indica un ús eficient de les dades per part del procés i un pes important dels sistemes d'informació com a base de la tramitació del procés.

Inferim a partir d'aquest indicador que les dades elementals que apareixen amb caràcter d'autòctones en una evidència tipus "dada" i es mantenen al llarg del procés en evidències tipus "dada" denoten un grau més alt de tramitació telemàtica o basada exclusivament en gestió de dades.

5.3.4.3 Índex de ròssec

De la mateixa manera que amb les sèries documentals, la importació de variables (o ròssec) es pot manifestar a dos nivells:

- Ròssec de variables entre les evidències d'un procés, que anomenarem **ròssec intraprocés**.
- Ròssec de variables entre les matrius de variables de processos relacionats, que anomenarem **ròssec interprocés**.

5.3.4.3.1 Ròssec intraprocés

Per obtenir l'índex es farà el mateix que amb les sèries documentals: dividirem el nombre d'ocurrències de cada dada que tenen la condició d'importades pel nombre teòric d'ocurrències importades que correspondrien a aquella dada. Per evitar distorsions en

l'indicador descartarem aquelles dades que només tenen una ocurrència, restant 1 al valor divisor.

PRO-17		
Indicador de procés	Índex de ròssec intraprocés	Interpretació de la mesura
Indicadors de base: PRO-08 Nombre d'ocurrències de dades RELO20-001 Tipus d'ocurrència	Fórmula de càlcul: ([PRO-08]Flt[RELO20-001]="IMP") / ([PRO-08]-1)	Un valor proper a 1 d'aquest índex indica una alta reutilització de cada dada i un desenvolupament òptim del procés o tramitació.

5.3.4.3.2 Ròssec interprocés

Com en les sèries documentals, aquest indicador es pot calcular de dues maneres. Proposarem la primera, que no té en compte el tipus d'ocurrència de la dada (autòctona o importada), sinó només la coincidències de la mateixa dada entre dos processos, preferentment relacionats. És per tant un indicador d'integració en potència, però no permet determinar si aquesta integració es materialitza. Per calcular aquest indicador, cal realitzar tres operacions:

1. A partir de la matriu de variables, obtenir el diccionari de cada procés.
2. A continuació recuperar les relacions entre processos recollides al mapa.
3. I en tercer lloc comparar els diccionaris de variables dels processos relacionats, i obtenir, a partir de les dades coincidents sobre el total de dades elementals d'ambdues sèries, l'índex de ròssec interprocés.

Aquest indicador, per tant, es calcula sempre sobre la intersecció de dos processos.

PRO-18		
Indicador de procés	Índex de ròssec interprocés	Interpretació de la mesura
Indicadors de base: PRO-07 Nombre de dades elementals	Fórmula de càlcul: [PRO-07] de [Procés 1] AND [Procés 2] / [PRO-07] de [Procés 1] OR [Procés 2]	Un valor elevat d'aquest indicador indica la possibilitat de reutilitzar dades entre dos processos, i en conseqüència de compartir o connectar els sistemes d'informació que els donen suport.

Considerem que el ròssec interprocés és un possible indicador de relació entre processos, i ho fem en base al supòsit que un procés conseqüent té:

- Una coincidència en les dades elementals amb el seu procés antecedent superior a la coincidència amb les d'altres processos no antecedents.
- En totes les variables coincidents, un major nombre de variables de ròssec o importades, i un menor nombre de variables autòctones.

5.3.5 Mapa de les relacions entre processos

Per a l'anàlisi visual de les relacions entre els processos es poden generar visualitzacions que mostrin de forma interactiva totes les relacions, i el tipus de relació. Per un exemple vegeu el cas d'aplicació al Consorci d'Educació de Barcelona a l'apartat 6.5.5 del capítol 6.

5.4 Indicadors sobre els sistemes d'informació

5.4.1 Mètriques descriptives dels sistemes d'informació

Les mètriques descriptives basades únicament en atributs dels sistemes d'informació són les següents:

SIS-01	
Indicador global	Nombre de sistemes d'informació
Nivells d'agregació	Per entorn
	Per tecnologia
	Per abast del sistema
	Per model de servei en la implantació
	Per nivell de seguretat (control d'accés + traçabilitat)

5.4.2 Mètriques de la transformació

La primera mètrica és correspon a l'evolució en el nombre de sistemes vigents en funció del seu entorn tecnològic.

SIS-02	
Indicador de sistema d'informació	Nombre de sistemes d'informació vigents
Nivells d'agregació	Per any i entorn

Podem obtenir-ne un major detall amb els següents indicadors. En primer lloc, el nombre de sistemes d'informació desenvolupats de nou cada any.

SIS-03	
Indicador de sistema d'informació	Nombre de sistemes d'informació desenvolupats
Nivells d'agregació	Per any i entorn

I en segon lloc, el nombre de sistemes d'informació que s'han deixat d'utilitzar cada any.

SIS-04	
Indicador de sistema d'informació	Nombre de sistemes d'informació desactivats
Nivells d'agregació	Per any i entorn

5.4.3 Mètriques derivades de la relació del sistema d'informació amb altres entitats

Les mètriques descriptives basades en les relacions dels sistemes d'informació amb altres entitats són les següents:

SIS-05	
Indicador de sistema d'informació	Nombre de processos
Nivells d'agregació	Per tipus de procés i entorn tecnològic

SIS-06	
Indicador de sistema d'informació	Nombre de sèries documentals
Nivells d'agregació	Per caràcter de la sèrie i entorn tecnològic

Els sistemes d'informació poden ser el repositori de les dades d'evidències d'una o més sèries documentals. El nombre de sèries (i de processos) als quals doni suport el sistema és una variable que pot determinar el seu caràcter (sistema transaccional o sistema gestor de dades mestres).

SIS-07	
Indicador de sistema d'informació	Nombre d'evidències
Nivells d'agregació	Per tecnologia

El nombre d'evidències que accedeixen a cada sistema d'informació es pot considerar un indicador rellevant del nivell d'utilització del sistema.

SIS-08	
Indicador de sistema d'informació	Nombre de relacions amb sistemes d'informació (nombre d'interfícies)
Nivells d'agregació	Per tipus d'interfície
	Per tipus d'importació
	Per procediment d'importació
	Per freqüència d'actualització
	Per tipus d'historificació

Aquest és un indicador bàsic per calcular les diferents modalitats de l'índex d'influència d'un sistema sobre altres sistemes d'informació (indicadors SIS-13, SIS-14, SIS-15 i SIS-16).

5.4.4 Mètriques sobre les dades del sistema d'informació

En aquest apartat els indicadors que inclouen la variable ocurrència de la dada es calculen tenint en compte les ocurrencies en evidències, mentre que els indicadors que només contempen la dada elemental com a unitat es calculen en base a les dades imputades directament al sistema.

D'acord amb això, el següent indicador compta el nombre de dades elementals assignades directament a un sistema d'informació. No és per tant directament equivalent als indicadors SER-06 i PRO-07, atès que un sistema pot contenir més dades elementals de les que apareixen en les evidències registrades. En cas de requerir aquesta equivalència, es podria calcular tanmateix amb el mateix criteri que es calculen els indicadors SER-06 i PRO-07, a partir de les evidències.

SIS-09	
Indicador de sistema d'informació	Nombre de dades elementals (assignades directament a sistemes d'informació)
Nivells d'agregació	Per tipus de dada
	Per objecte
	Per superobjecte
	Per tipus d'objecte (cosa/acció)

La segona forma de vincular dades a sistemes és a través de les variables que apareixen en evidències del tipus "Dades", vinculades sempre a un sistema d'informació. Això ens permet incorporar en els sistemes la variable "Caràcter" de l'ocurrència, per indicar si és autòctona o importada. Aquest indicador sí que és equivalent als indicadors SER-07 i PRO-08.

SIS-10	
Indicador de sistema d'informació	Nombre d'ocurrències de dades (dades en evidències assignades al sistema)
Nivells d'agregació	Per tipus d'ocurrència (autòctona o importada)

Es planteja també, amb els indicadors SIS-11 i SIS-12, obtenir mètriques a nivell d'agregacions de dades en objectes i superobjectes.

SIS-11		
Indicador de sistema d'informació	Nombre d'objectes	Interpretació de la mesura
Nivells d'agregació	Per tipus d'objecte (cosa/acció)	Base per al càlcul de l'índex de transaccionalitat basat en objectes (SIS-18).

SIS-12		
Indicador de sistema d'informació	Nombre de superobjectes	Interpretació de la mesura
Nivells d'agregació	Per tipus de superobjecte (cosa/acció)	

Aquests són dos indicadors realment significatius del caràcter d'un sistema d'informació, i esdevenen la base per al càlcul de l'índex de transaccionalitat basat en objectes (indicador SIS-18).

5.4.5 Indicadors d'anàlisi dels sistemes d'informació

5.4.5.1 Índex d'influència del sistema d'informació

Aquest índex pretén mesurar la influència que pot exercir un sistema d'informació sobre altres sistemes de l'organització, i al mateix temps fins a quin punt n'és dependent. L'indicador es planteja en quatre modalitats.

5.4.5.1.1 Índex de criticitat

SIS-13		
Indicador de sistema d'informació	Índex de criticitat (índex d'influència simple)	Interpretació de la mesura
Indicadors de base: SIS-08 Nombre de relacions amb sistemes d'informació (nombre d'interfícies)	Fórmula de càlcul: [SIS-08]	Un valor elevat d'aquest indicador revela una alta participació del sistema en intercanvis d'informació.

Aquest indicador equival de forma directa a l'indicador SIS-08. La lectura a nivell de sistema permet assenyalar determinats sistemes com a crítics. La rellevància real del sistema en aquests intercanvis de dades ens la donarà l'índex d'influència (SIS-14), però aquest indicador ja ens permet, pels valors més alts, assenyalar els possibles sistemes crítics.

5.4.5.1.2 Índex d'influència

SIS-14		
Indicador de sistema d'informació	Índex d'influència	Interpretació de la mesura
Indicadors de base: SIS-08 Nombre de relacions amb sistemes d'informació (nombre d'interfícies) ENT011-001 Identificador (de sistema d'informació) ENT012 Interfície de connexió entre sistemes	Fórmula de càlcul: $([SIS-08]Flt[ENT011-001][ENT012]="Origen") - ([SIS-08]Flt[ENT011-001][ENT012]="Destí")$	Un sistema amb un índex d'influència positiu serà un sistema proveïdor de dades. Per altra banda, un sistema amb un alt nombre de sistemes proveïdors (indicador negatiu) pot indicar que es tracta d'un sistema orientat al consum de la informació (explotació) o bé a la consolidació i arxiu.

Aquest indicador es calcula a partir de la diferència entre:

- La quantitat d'interaccions per sistema i tipus d'interacció (sistemes origen). És a dir, número de sistemes als quals el sistema facilita dades. Com més alt sigui aquest número, més influent es considerarà aquest sistema.
- La quantitat d'interaccions per sistema i tipus d'interacció (sistemes destí). És a dir, número de sistemes que proveeixen dades al sistema. Com més alt sigui aquest número, més dependent es considerarà aquest sistema d'altres sistemes.

Aquest indicador defineix el rol d'un sistema en relació amb la resta de sistemes de l'organització. De totes maneres, per validar realment el caràcter d'un sistema com a proveïdor de dades considerem que caldria incorporar una segona variable, que és el nombre de dades autòctones que conté el sistema, és a dir el nombre de dades que es recullen per primer cop en aquest sistema.

Un exemple de la distorsió que pot provocar no considerar aquesta variable és l'aparició de SIDIE com un dels sistemes més influents del Consorci. SIDIE, que realment aporta dades a nombrosos sistemes del Consorci, és un sistema de *Business Intelligence* que alimenta el Departament d'Educació, i que obté les seves dades de nombrosos sistemes d'informació del Departament. Per tant, malgrat el rol que ocupa en el Consorci és de sistema proveïdor d'informació, en el context tecnològic del Departament d'Educació SIDIE és un sistema altament dependent d'altres sistemes.

5.4.5.1.3 Índex de connectivitat ponderada

SIS-15		
Indicador de sistema d'informació	Índex de connectivitat ponderada	Interpretació de la mesura
Indicadors de base: SIS-08 Nombre de relacions amb sistemes d'informació (nombre d'interfícies) ENT012-005 Procediment d'importació	Fórmula de càlcul: [SIS-08]*([Pes]Fit[ENT012-005])	Un valor alt d'aquest indicador indica una major facilitat del sistema per connectar-se amb altres sistemes, i especialment de forma automatitzada.

Aquest indicador mesura quan de connectable o obert és el sistema, en funció del nombre i tipus dels intercanvis que tingui amb altres sistemes. En aquest cas no es distingeix entre sistema proveïdor i receptor, i es compten totes les interaccions o interfícies sense distinció. Es considera que com més automatitzada sigui la connexió (i per tant amb un valor més alt d'aquest índex), més connectable (obert) és el sistema.

Per ponderar el tipus de connexió, s'aplica a cada interacció la següent taula de valors:

Taula 5-13: Pes en funció del tipus de connexió entre sistemes d'informació

Tipus de connexió	Pes
Connexió directa a origen	3
Robotitzat	2
Manual directe	1
Manual amb depuració prèvia	0

5.4.5.1.4 Índex de connectivitat potencial

L'existència d'una alta coincidència en les dades de dos sistemes d'informació és un indicador que permet quantificar els beneficis d'establir una connexió, preferiblement automatitzada, entre aquells dos sistemes, atès que la coincidència en el model de dades evitarà realitzar canvis en els sistemes que intercanvien les dades.

SIS-16		
Indicador de sistema d'informació	Índex de connectivitat potencial	Interpretació de la mesura
Indicadors de base: SIS-09 Nombre de dades elementals (assignades directament a sistemes d'informació)	Fórmula de càlcul: [SIS-09] de [Sistema 1] AND [Sistema 2] / [SIS-09] de [Sistema 1] OR [Sistema 2]	Un índex alt de coincidència entre els diccionaris de dades de dos sistemes d'informació indica un potencial...

SIS-16		
Indicador de sistema d'informació	Índex de connectivitat potencial	Interpretació de la mesura
		...d'intercanvi d'informació que pot estalviar en molt la recollida de dades a un dels dos sistemes.

Posem un exemple de l'ús d'aquest indicador utilitzant dades del Consorci d'Educació de Barcelona. Si estem valorant el potencial d'integració del sistema SI04.050 Comunicació maltractament tipus lleu amb altres sistemes del Consorci, observem que aquest sistema comparteix dades amb aquests altres sistemes del Consorci (índex superior a 40):

Taula 5-14: Índex de connectivitat potencial del sistema d'informació SI04.053

SISTEMA ORIGEN	SISTEMA DESTÍ	ÍNDEX
SI04.050 Comunicació maltractament tipus lleu	SI04.049 Comunicació maltractament tipus greu	71,43
	SI01.078 Registre maltractaments i assetjaments	60,00
	SI01.127 Registre d'alumnes en atenció domiciliària o hospitalària	57,14
	SI01.032 Registre de compatibilitats	50,00
	SI01.144 Eina de registre i seguiment dels alumnes del Èxit 2 - C	50,00
	SI01.072 Registre de resolucions	42,86
	SI04.014 Eleccions als Consells Escolars (RECE)	42,86
	SI01.005 BBDD de l'alumnat estranger	40,00

El major índex de coincidència el trobem amb el SI04.049, cosa la qual podria indicar que es podria valorar la fusió en un sol sistema. En els dos casos següents, SI01.078 i SI01.127, podríem valorar possibilitats de connectivitat. En aquest cas, caldria valorar aleshores la justificació funcional i d'esforç, i les possibilitats tecnològiques de fer efectiva la connexió.

Aquest indicador s'ha de complementar però amb el context funcional de cada sistema, atès que la mera coincidència, basada realment només en el contingut, no implica que tingui sentit, des d'un punt de vista operatiu, l'intercanvi de dades entre els dos sistemes, com s'esdevé probablement amb els sistemes de l'exemple (SI01.032 i SI04.014).

5.4.5.2 Índex de transaccionalitat del sistema

De la mateixa manera que amb les sèries documentals, l'índex de transaccionalitat vol recollir el volum de contingut del sistema d'informació que té caràcter o significació transaccional. Aquest indicador té tres variants, que es detallen a continuació.

5.4.5.2.1 Índex de transaccionalitat calculat a partir de l'objecte de les dades

L'objectiu d'aquesta operació és obtenir un indicador que representi el caràcter majoritari de les dades del sistema. Aquest indicador es calcularà només sobre les dades elementals i els seus objectes assignades directament al sistema, sense tenir en compte les diferents ocurrencies de cada dada en diferents evidències associades al sistema. Consisteix en el percentatge de dades elementals del sistema l'objecte de les quals és equiparable a acció sobre el conjunt de dades del sistema.

SIS-17		
Indicador de sistema d'informació	Índex de transaccionalitat a partir de les dades elementals	Interpretació de la mesura
Indicadors de base: SIS-09 Nombre de dades elementals per sistema d'informació i tipus d'objecte (cosa/acció) ENT014-002 Tipus d'objecte	Fórmula de càlcul: ([SIS-09]Flt[ENT014-002]="Acció") / [SIS-09]	La hipòtesi que s'utilitzarà per categoritzar el sistema d'informació com a transaccional serà la següent: es considerarà que un sistema amb un índex de transaccionalitat basat en dades superior a 0,5 (50%) té més probabilitats de ser un sistema transaccional que un amb un índex inferior. ²³⁵

5.4.5.2.2 Índex de transaccionalitat calculat a partir del tipus de dades

Existeix una segona forma de calcular aquest índex, menys precisa, però també més ràpida. Es basa en el tipus de dada, i en el recompte del nombre de dades de tipus transaccional sobre el total de dades del sistema d'informació. La fórmula de càlcul és la següent:

SIS-18		
Indicador de sistema d'informació	Índex de transaccionalitat a partir del tipus de dades	Interpretació de la mesura
Indicadors de base: SIS-09 Nombre de dades elementals per sistema d'informació	Fórmula de càlcul: ([SIS-09]Flt[ENT006-003]="Transaccional") / [SIS-09]	Un valor alt d'aquest índex indica que el sistema recull un nombre elevat d'informació sobre...

²³⁵ Aquesta hipòtesi es basa en l'observació que els sistemes transaccionals habitualment només recullen les dades mínimes imprescindibles dels objectes tipus "cosa" per a fer-hi referència quan hi associen les dades de tramitació o transaccionals, i no tenen per objectiu recollir descripcions àmplies de coses. Hipòtesi que intentem verificar amb l'objectiu 4 (apartat 8.1.4 del capítol 8).

SIS-18		
Indicador de sistema d'informació	Índex de transaccionalitat a partir del tipus de dades	Interpretació de la mesura
ENT006-003 Tipus de dada		...l'execució dels processos (dates de fites, canvis d'estat, etc.) que identifica aquest sistema com a sistema de gestió d'aquell procés, i per tant de caràcter transaccional.

Un objecte del tipus “acció” no es descriu només amb dades transaccionals, sinó també amb dades identificatives, descriptives, etc. La transaccionalitat d'un sistema es mesura doncs de forma més precisa amb l'índex SIS-17 que no amb l'índex SIS-18, que pot donar falsos negatius en aquells casos on cap de les dades que descriuen un objecte “acció” és de tipus transaccional. Vegem alguns exemples:

Taula 5-15: Índex de transaccionalitat a partir del tipus de dades d'alguns sistemes d'informació del Consorci

SISTEMA	TOTAL DADES	DADES ACCIO	DADES TRANSACCIONALS	ÍNDEX SIS-17	ÍNDEX SIS-18
SI01.191 Eina de seguiment de la contractació de tècnics	4	4	0	1,00	0,00
SI01.193 Eina seguiment estat de la licitació	11	11	0	1,00	0,00
SI01.029 Registre del Pressupost	9	9	1	1,00	0,11
SI08.014 DOC003-001-L1 Transferències a Arxiu Central	3	3	1	1,00	0,33
SI08.020 SGRH020-005 Formació del personal	4	4	1	1,00	0,25
SI08.024 SLOP001-001 Desenvolupament del sistema de protecció de dades [...]	4	4	3	1,00	0,75

Analizant una mostra aleatòria de sis sistemes d'informació del Consorci, observem que la manca de cap dada etiquetada com a dada transaccional en els sistemes SI01.191 Eina de seguiment de la contractació de tècnics i SI01.193 Eina seguiment estat de la licitació provoca dos falsos negatius. En els altres quatre casos, veiem com sistemes que només descriuen objectes de tipus “acció” obtenen un índex de transaccionalitat baix a causa del baix nombre de dades classificades com transaccionals.

Aquest és un risc que hem observat que té una freqüència elevada en el conjunt de l'indicador, i que per tant desaconsella l'ús independent de l'indicador SIS-18 per a mesurar la transaccionalitat d'un sistema d'informació. Això no obstant, aquells sistemes que obtenen un fals negatiu o un índex extremadament baix de dades transaccionals és aconsellable que siguin revisats per verificar que no es tracta d'un error d'atribució, o que realment són registres d'accions que no s'utilitzen per a cap tasca de gestió.

5.4.5.2.3 Índex de transaccionalitat calculat a partir dels objectes

Un sistema pot contenir dades descriptives de diferents objectes, siguin coses o accions. L'índex de transaccionalitat basat en els objectes descrits pel sistema es calcula considerant els objectes descrits pel sistema en relació al total d'objectes dels quals aquell sistema gestiona dades.

SIS-19		
Indicador de sistema d'informació	Índex de transaccionalitat a partir dels objectes	Interpretació de la mesura
Indicadors de base: SIS-11 Nombre d'objectes per sistema d'informació ENT014-002 Tipus d'objecte	Fórmula de càlcul: ([SIS-11]Fit[ENT014-002]="Acció") / [SIS-11]	Es considera que un sistema amb un alt índex de transaccionalitat basat en objectes té més probabilitats de ser un sistema transaccional (o de gestió), atès que els sistemes transaccionals fan referència a més objectes (una acció i diverses coses que intervenen en l'acció) que no pas un sistema de dades mestres. ²³⁶

5.4.5.3 Classificació dels sistemes pel nivell de seguretat de les dades

La darrera mètrica consisteix en obtenir una mesura representativa del nivell de seguretat de les dades contingudes en un sistema d'informació.

²³⁶ Hipòtesi que s'ha intentat verificar amb l'objectiu 4 (apartat 8.1.4 del capítol 8).

SIS-20		
Indicador de sistema d'informació	Classificació per nivell de seguretat de les dades	Interpretació de la mesura
Indicadors de base: Atribut ENT011-008 Entorn tecnològic Atribut ENT011-010 Abast Atribut ENT011-013 Controls d'integritat Atribut ENT011-014 Controls d'accés	Fórmula de càlcul: $([Pes]Flt[ENT011-008]) + ([Pes]Flt[ENT011-010]) + [ENT011-013] + [ENT011-014]$	El valor de l'indicador determina el nivell de fiabilitat imputable a les dades en funció dels sistemes d'autenticació i traçabilitat inferits dels atributs components de l'indicador.

Aquesta classificació s'adreça a agrupar els sistemes d'informació en funció de la capacitat per garantir la integritat i l'autenticitat de les dades contingudes. L'assignació de la categoria es fa a partir de la ponderació d'una sèrie de característiques del sistema d'informació, les quals determinen el nivell de seguretat i els mecanismes per garantir la integritat de les dades.

Les característiques que determinen la ponderació són les següents:

- La tecnologia de base del sistema, recollida a l'atribut **ENT011-008 Entorn tecnològic**. S'ha aplicat la següent ponderació:

Taula 5-16: Pes en funció de l'entorn tecnològic del sistema d'informació

Entorn tecnològic ²³⁷	Pes
Sistemes d'informació en entorns corporatius i <i>mainframe</i>	4
Desenvolupaments sobre tecnologia professional	3
Solucions <i>cloud (low code)</i> i formularis	2
Solucions d'oficina	1

- L'abast, recollit a l'atribut **ENT011-010 Abast**, el qual determina l'existència d'un departament de tecnologies de la informació que pugui garantir uns determinats nivells de seguretat.

Taula 5-17: Pes en funció de l'abast del sistema d'informació

Abast	Pes
Sistemes corporatius d'Administracions públiques	2
Desenvolupaments propis del CEB o d'empresa externa	1
Solucions d'usuari final	0

²³⁷ L'impacte en la seguretat d'aquest atribut ve determinat pel tipus de professionals que pot desenvolupar sistemes. Així, mentre que els sistemes en entorns corporatius i *mainframe* han de ser desenvolupats necessàriament per professionals informàtics, les solucions *low code* poden ser desenvolupades per usuaris avançats, mentre que les solucions d'oficina poden ser posades en marxa per usuaris sense cap formació tecnològica específica.

- L'existència de controls d'integritat o traçabilitat (atribut **ENT011-013 Controls d'integritat**), i de controls d'accés i ús (atribut **ENT011-014 Controls d'accés**).

S'ha decidit no utilitzar la classificació d'acord amb l'Esquema Nacional de Seguridad²³⁸ (atribut **ENT011-015 Classificació ENS**) per dos motius. En primer lloc, perquè en el marc de la Generalitat de Catalunya i de l'Ajuntament de Barcelona només s'aplica a un conjunt de sistemes crítics, que coincideix amb un percentatge molt baix dels sistemes analitzats en aquest projecte. I en segon lloc, perquè no ha estat accessible durant el treball de camp per a sistemes analitzats que sí que disposen d'aquesta classificació.

5.4.6 Mapa de relacions entre sistemes d'informació

Per a l'anàlisi visual de les relacions entre sistemes d'informació es poden generar visualitzacions que mostrin de forma interactiva les interaccions entre sistemes, i el tipus d'interacció. Per un exemple vegeu el cas d'aplicació al Consorci d'Educació de Barcelona a l'apartat 6.6.6 del capítol 6.

²³⁸ El marc regulatori de la seguretat informàtica al sector públic català ve determinat fonamentalment pel *Real Decreto 311/2022, de 3 de mayo, por el que se regula el Esquema Nacional de Seguridad*, i per les pautes i recomanacions de l'Agència de Ciberseguretat de Catalunya (<https://ciberseguretat.gencat.cat/ca/inici/index.html>).

5.5 Indicadors sobre les dades

5.5.1 Mètriques descriptives de les dades elementals

Les mètriques descriptives basades únicament en atributs de les dades són les següents:

DAD-01	
Indicador global	Nombre de dades elementals
Nivells d'agregació	Per tipus de dada
	Per objecte
	Per superobjecte
	Per tipus d'objecte (cosa/acció)

Aquest indicador és la base per al càlcul de l'índex de rellevància de la dada elemental (indicador DAD-23).

El segon indicador és el nombre d'ocurrències, per al qual s'indiquen els nivells d'agregació més habituals.

DAD-02	
Indicador global	Nombre d'ocurrències de dades
Nivells d'agregació	Per dada elemental
	Per tipus de dada
	Per objecte
	Per superobjecte
	Per tipus d'objecte (cosa/acció)
	Per tipus d'ocurrència (autòctona/importada)
	Per entorn del sistema
	Per tipus d'evidència

Cal tenir present que, tal com ja s'ha comentat, es disposa de dues fonts per a les ocurrències: les assignacions directes a sistemes, i les assignacions a evidències. Per tant, l'indicador DAD-02 es pot calcular utilitzant aquestes dues fonts en funció de amb quina altra mesura s'hagi de comparar.

Per a la descripció dels nivells agregats de les dades elementals es proposen dos indicadors.

DAD-03	
Indicador global	Nombre d'objectes
Nivells d'agregació	Per tipus de dada
	Per superobjecte
	Per tipus d'objecte (cosa/acció)

DAD-04	
Indicador global	Nombre de superobjectes
Nivells d'agregació	Per tipus de dada
	Per tipus d'objecte (cosa/acció)

5.5.2 Mètriques derivades de la relació de la dada elemental amb altres entitats

Es proposen a continuació 16 mètriques descriptives derivades de les relacions de les dades amb altres entitats. Aquestes mètriques es distribueixen en funció dels quatre nivells d'agrupació de les dades (dades elementals, objectes, superobjectes i tipus d'objecte).

5.5.2.1.1 A nivell de dada elemental

DAD-05	
Indicador de dada elemental	Nombre de sistemes d'informació
Nivells d'agregació	Per entorn/abast del sistema

DAD-06	
Indicador de dada elemental	Nombre d'evidències
Nivells d'agregació	Per tipus d'evidència

DAD-07	
Indicador de dada elemental	Nombre de processos
Nivells d'agregació	Per tipus de procés

DAD-08	
Indicador de dada elemental	Nombre de sèries documentals
Nivells d'agregació	Per caràcter de la sèrie

Aquests quatre indicadors són la base per al càlcul de l'índex de rellevància global de la dada elemental, recollit en l'indicador DAD-23.

5.5.2.1.2 A nivell d'objecte

DAD-09	
Indicador d'objecte	Nombre de sistemes d'informació
Nivells d'agregació	Per entorn/abast del sistema

DAD-10	
Indicador d'objecte	Nombre d'evidències
Nivells d'agregació	Per tipus d'evidència

DAD-11	
Indicador d'objecte	Nombre de processos
Nivells d'agregació	Per tipus de procés

DAD-12	
Indicador d'objecte	Nombre de sèries documentals
Nivells d'agregació	Per caràcter de la sèrie

Aquests quatre indicadors són la base per al càlcul de l'índex de rellevància global de l'objecte, recollit en l'indicador DAD-24.

5.5.2.1.3 A nivell de superobjecte

DAD-13	
Indicador de superobjecte	Nombre de sistemes d'informació
Nivells d'agregació	Per entorn/abast del sistema

DAD-14	
Indicador de superobjecte	Nombre d'evidències
Nivells d'agregació	Per tipus d'evidència

DAD-15	
Indicador de superobjecte	Nombre de processos
Nivells d'agregació	Per tipus de procés

DAD-16	
Indicador de superobjecte	Nombre de sèries documentals
Nivells d'agregació	Per caràcter de la sèrie

Aquests quatre indicadors són la base per al càlcul de l'índex de rellevància global del superobjecte, recollit en l'indicador DAD-25.

5.5.2.1.4 A nivell de tipus d'objecte (cosa/acció)

DAD-17	
Indicador de tipus d'objecte	Nombre de sistemes d'informació
Nivells d'agregació	Per entorn/abast del sistema

DAD-18	
Indicador de tipus d'objecte	Nombre d'evidències
Nivells d'agregació	Per tipus d'evidència

DAD-19	
Indicador de tipus d'objecte	Nombre de processos
Nivells d'agregació	Per tipus de procés

DAD-20	
Indicador de tipus d'objecte	Nombre de sèries documentals
Nivells d'agregació	Per caràcter de la sèrie

5.5.3 Indicadors d'anàlisi de les dades elementals

5.5.3.1 Índex de rellevància de la dada elemental

5.5.3.1.1 Índex de rellevància de la dada elemental dins el procés o la sèrie documental

Tot procés, i tota sèrie documental, tenen un conjunt de dades clau. Per identificar aquest conjunt es proposa calcular la rellevància de la dada elemental dins el procés o la sèrie documental. Per a realitzar aquest càlcul es tindran en compte dues variables:

1. El nombre d'ocurrències de la dada elemental dins el procés o la sèrie documental.
2. El fet que la dada s'origini o reculli per primer cop en aquell procés o sèrie (condició d'autòctona).

Aquest indicador dona com a resultat una matriu de procés o sèrie, on cada dada té un pes determinat en funció de quantes vegades apareix com autòctona o importada.

DAD-21		
Indicador de dada elemental	Índex de rellevància dins el procés o la sèrie documental	Interpretació de la mesura
Indicadors de base: DAD-02 Nombre d'ocurrències de dades RELO20-001 Tipus d'ocurrència	Fórmula de càlcul: [DAD-02] + ([DAD-2] Flt[RELO20-001]="AUT")	Les dades que obtenen la magnitud més alta en aquest indicador constitueixen les dades clau o essencials del procés o sèrie documental, i les que hauran de ser objecte d'especial protecció i, si escau, preservació.

En el cas per exemple de la sèrie documental INF001-003 Elaboració d'estudis amb col·laboracions externes, del Consorci d'Educació de Barcelona, les dades amb un índex de rellevància superior són:²³⁹

Taula 5-18: Índex de rellevància de les dades elementals de la sèrie documental INF001-003

DADA ELEMENTAL	ÍNDEX
Nom unitat administrativa	6,00
Observacions - informe	6,00
Data de l'informe	4,00
Nom de l'informe	4,00
Nom i cognoms treballador CEB	4,00
Servei o projecte	4,00
Correu electrònic treballador CEB	3,00
Nom i cognoms persona externa	3,00
Revisat	3,00
Abast geogràfic	2,00
Correu corporatiu	2,00
Correu electrònic entitat	2,00
Descriptors estudis	2,00
Etape educativa	2,00
NIF entitat	2,00
Nom entitat col·laboradora	2,00
Publicat Web	2,00

Aquest indicador té dos punts febles. El primer punt és que sobrerrepresenta a les dades menys integrades, aquelles que cal informar manualment més vegades. Un procés poc integrat incrementarà doncs l'índex de rellevància de la dada, sense que això impliqui una importància real d'aquella dada, més enllà del seu pes en l'operativa del procés. I el segon punt feble és que requereix que les dades que ja s'han recollit en processos anteriors es marquin com importades en la primera evidència del procés conseqüent.

²³⁹ Vegeu taula 5-5 per a la matriu de variables origen d'aquests valors (apartat 5.2.4.3.1 d'aquest capítol).

Atès que aquesta no és la pauta que s'ha seguit en el cas d'aplicació del Consorci, on totes les primeres aparicions de cada dada elemental en un procés o sèrie s'han marcat com autòctones, no podem per tant posar a prova aquest indicador amb el cas del Consorci. Considerem doncs que aquest indicador ha de restar en qualitat de proposta, pendent d'una anàlisi més detallada de les seves casuístiques.

5.5.3.1.2 Índex de rellevància de la dada elemental dins el sistema d'informació

Aquest indicador és una variant de l'anterior. Les dades elementals només existeixen una vegada dins de cada sistema d'informació. Per tant, per calcular la rellevància de la dada elemental a nivell de sistema cal treballar amb les assignacions de dades elementals a evidències, que és la variable a través de la qual podem detectar l'ús de la dada elemental. Treballarem novament amb dues variables:

1. El nombre d'ocurrències de la dada elemental del sistema dins de diferents processos o sèries documentals de l'organització (nombre d'evidències on apareix aquella dada). Tindrem en compte només les evidències basades en dades (les vinculades directament a aquell sistema), atès que la rellevància global de cada dada ja l'obtidrem amb el càlcul de l'índex anterior.
2. El fet que la dada aparegui en l'evidència amb la condició d'autòctona.

DAD-22		
Indicador de dada elemental	Índex de rellevància dins el sistema d'informació	Interpretació de la mesura
Indicadors de base: DAD-02 Nombre d'ocurrències de dades en totes les evidències del tipus "dades" d'un sistema d'informació RELO20-001 Tipus d'ocurrència ENT004-003 Tipus d'evidència	Fórmula de càlcul: $[DAD-02] + ([DAD-2] \text{Flt}[RELO20-001]="AUT" \text{ AND } [ENT004-003]="Dades")$	Com en l'indicador anterior, una mesura alta d'aquest indicador ens assenyala el conjunt de dades clau d'un sistema d'informació pel que fa a la seva freqüència d'ús i originalitat.

Continuant amb l'exemple anterior, el sistema d'informació que dona suport a la sèrie documental INF001-003 Elaboració d'estudis amb col·laboracions externes és el sistema SI08.023 SIN001-007 Funcionament del Gabinet d'estudis del CEB. L'índex calculat sobre les dades d'aquest sistema dona el següent resultat:

Taula 5-19: Índex de rellevància de les dades elementals del sistema d'informació SI08.023

DADA ELEMENTAL	ÍNDEX
Abast geogràfic	2,00
Data de l'informe	2,00
Descriptors estudis	2,00
Etape educativa	2,00
Nom de l'informe	2,00
Nom unitat administrativa	2,00
Observacions - informe	2,00
Publicat Web	2,00
Revisat	2,00
Servei o projecte	2,00
Nom i cognoms treballador CEB	1,00

Novament aquest indicador té el punt feble que sobrerrepresenta a les dades menys integrades, aquelles que cal informar manualment més vegades, tot i que, considerant que ho valorem dins un mateix sistema d'informació, aquest fet és molt menys freqüent que quan ho valorem des de la perspectiva de procés o de sèrie documental.

5.5.3.1.3 Índex de rellevància global de la dada elemental (dins el mapa de dades)

En els dos indicadors anteriors hem utilitzat únicament el nombre d'ocurrències per mesurar la rellevància de la dada. Aquesta tercera mesura pretén objectivar la rellevància de la dada dins el conjunt del mapa de dades, en funció del nombre d'entitats o actius d'informació als quals afecta. La rellevància global vindrà determinada pel nombre d'usos de la dada vinculats a diferents entitats: processos, sèries documentals, evidències i sistemes d'informació.

En aquest càlcul cal ponderar separatament les ocurrències de la dada a nivell de sistema, i les ocurrències consignades dins d'evidències, atès que totes les ocurrències en evidències basades en dades coincidirán amb dades imputades directament al sistema d'informació. No totes les dades del sistema apareixerán en el procés, però totes les dades d'evidències basades en dades també estaran informades en algun sistema.

Per donar el pes adient a l'assignació, hem de diferenciar quines assignacions generen el valor original, o primera aparició de la dada a l'organització (variables autòctones en l'assignació), i quines assignacions generen un valor ja existent prèviament a l'organització (variables importades o de ròssec). D'aquesta manera:

- Una variable pot ser autòctona per a un procés (la primera vegada que apareix en aquell procés) però heretada per al conjunt de l'organització (haver aparegut com autòctona en processos anteriors al procés analitzat).
- Dues variables poden ser autòctones per al conjunt de l'organització, si els mecanismes de captura no estan integrats.

Aquests usos es recolliran utilitzant les següents mesures:

- Nombre de processos (vigents) on apareix la dada elemental
- Nombre de sèries documentals (vigents) on apareix la dada elemental
- Nombre d'evidències (vigents) on apareix la dada elemental
- Nombre de sistemes d'informació (vigents) on apareix la dada elemental
- Nombre de visualitzacions de dades (vigents) on s'utilitza la dada elemental²⁴⁰

La suma dels cinc indicadors dividida per la meitat del total d'actius d'informació considerats serà l'índex de rellevància global sobre el conjunt del mapa de dades.²⁴¹

Però probablement aquest pes no és suficientment representatiu de la rellevància real de la dada. Per aquest motiu, s'ha valorat ponderar aquest pes atribuït a cada entitat un pes basat en alguns dels indicadors proposats anteriorment:

- Nombre de processos (vigents) on apareix la dada elemental on cada procés suma el seu índex d'influència
- Nombre de sèries documentals (vigents) on apareix la dada elemental on cada sèrie suma el seu índex d'influència
- Nombre d'evidències (vigents) on apareix la dada elemental on la dada suma el seu caràcter d'autòctona
- Nombre de sistemes d'informació (vigents) on apareix la dada elemental on cada sistema suma el seu índex d'influència

²⁴⁰ Com a part del projecte també s'ha començat a recollir informació sobre les dades elementals utilitzades en les visualitzacions de dades i els quadres de comandament. Aquesta recollida no està finalitzada, i no proporciona per tant dades suficients per incorporar-la en aquesta anàlisi. Això no obstant, s'inclou com un dels ítems que formaran part en un futur d'aquest indicador.

²⁴¹ Per l'experiència del cas d'aplicació, considerem que és altament improbable que una dada aparegui en més de la meitat dels actius d'informació d'una organització, motiu pel qual fixem el valor 1 d'aquest indicador en l'aparició de la dada en el 50% d'actius o entitats considerades.

Considerem en aquesta proposta només la primera variant de càlcul d'aquest indicador.

DAD-23		
Indicador de dada elemental	Índex de rellevància global	Interpretació de la mesura
Indicadors de base: DAD-07 Nombre de processos DAD-08 Nombre de sèries documentals DAD-06 Nombre d'evidències DAD-05 Nombre de sistemes d'informació ENT006-001 Identificador	Fórmula de càlcul: $\frac{([DAD-07]+[DAD-08]+[DAD-06]+[DAD-05])Fit[ENT006-001]}{([DAD-07]+[DAD-08]+[DAD-06]+[DAD-05])/2}$	Les dades amb un índex absolut de rellevància més gran es pot inferir que són dades més utilitzades, amb més presència en més entorns. ²⁴²

5.5.3.2 Índex de rellevància de l'objecte i del superobjecte

5.5.3.2.1 Índex de rellevància global de l'objecte i del superobjecte

La forma simple de calcular la rellevància de l'objecte és sumar el nombre de dades elementals que descriuen cada objecte i cada superobjecte. Aquesta mesura és la que es recull amb l'indicador DAD-01 Nombre de dades elementals agregat a nivell d'objecte o superobjecte.

Però podem aplicar el mateix raonament que en l'indicador DAD-23, i considerar que la rellevància de l'objecte vindrà donada per la suma de:

- Nombre de processos (vigents) que actuen sobre l'objecte/superobjecte
- Nombre de sèries documentals (vigents) que documenten l'objecte/superobjecte
- Nombre d'evidències (vigents) que fan referència a l'objecte/superobjecte
- Nombre de sistemes d'informació (vigents) que tenen dades sobre l'objecte/superobjecte
- Nombre de visualitzacions de dades (vigents) on es visualitzen dades de l'objecte/superobjecte

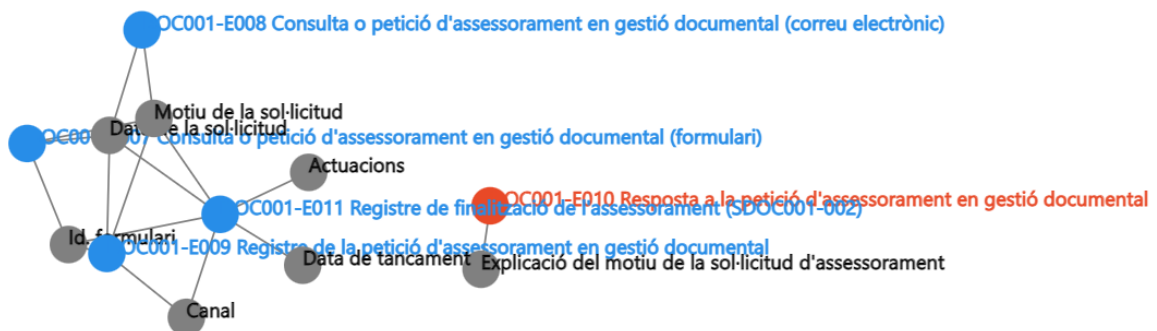
²⁴² Una possible conclusió és que són dades candidates a dades mestres, i per extensió els objectes que descriuen també són candidats a objectes de taules mestres. Això no és extrapolable directament als sistemes, ja que un sistema pot utilitzar una dada molt freqüent i fer-ho en un entorn completament transaccional. Per tant, per a l'extrapolació als sistemes tindrem en compte l'objecte majoritari del sistema, i analitzarem si aquest objecte coincideix amb l'objecte de la dada "mestra".

Aquest és el que podem anomenar **índex de rellevància global de l'objecte**. La fórmula de càlcul en funció del nivell d'agregació serà la següent:

DAD-24		
Indicador d'objecte	Índex de rellevància global	Interpretació de la mesura
Indicadors de base: DAD-11 Nombre de processos DAD-12 Nombre de sèries documentals DAD-10 Nombre d'evidències DAD-09 Nombre de sistemes d'informació	Fórmula de càlcul: $\frac{([DAD-11]+[DAD-12]+[DAD-10]+[DAD-09]) \cdot \text{Flt}[ENT006-001]}{([DAD-11]+[DAD-12]+[DAD-10]+[DAD-09])/2}$	Un valor alt en aquest índex indica una major presència de l'objecte en un conjunt d'activitats, els documents que se'n generen i els sistemes d'informació que hi donen suport.

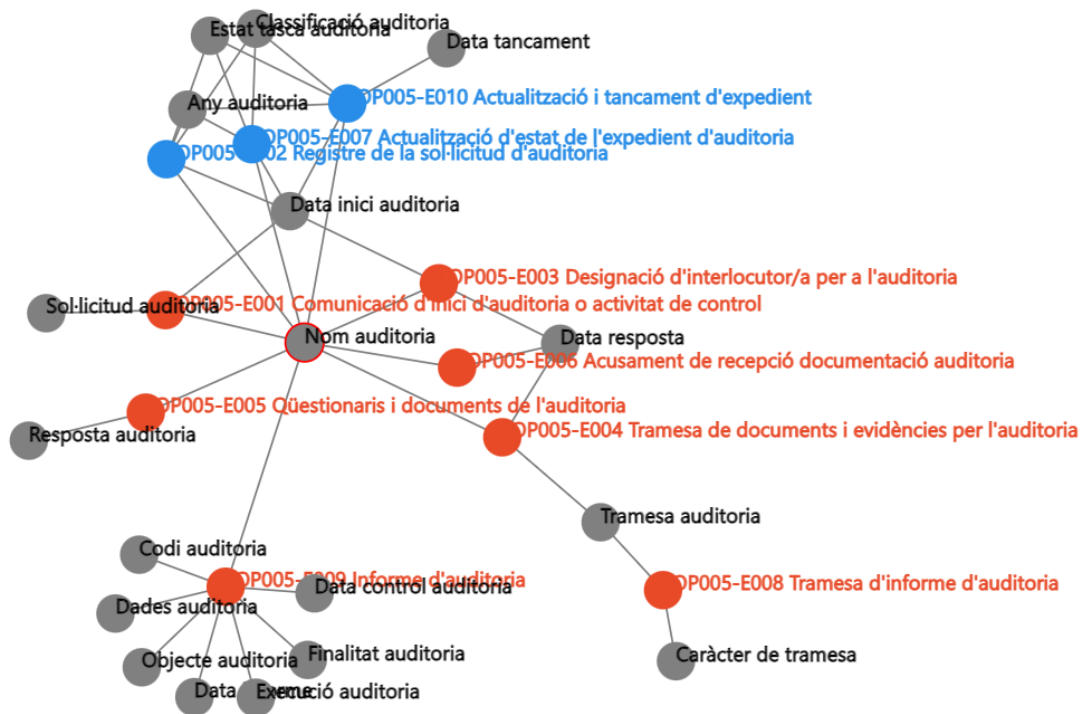
Aquest índex també pot tenir una representació visual, il·lustrativa de l'activitat que es genera al voltant de cada objecte de gestió. Agafant un exemple del Consorci d'Educació de Barcelona (limitat a evidències i dades) comparem dos objectes transaccionals. El primer, la gestió d'un assessorament a un centre educatiu en matèria de gestió documental, genera aquesta activitat, en forma de dades (en gris), evidències-objecte (en vermell) i interaccions amb sistemes d'informació (en blau):

Figura 5-9: Mapa de dades elementals, evidències i sistemes d'informació de l'objecte "Assessorament a un centre educatiu en matèria de gestió documental"



El segon, l'atenció a una auditoria de protecció de dades personals, genera aquesta activitat:

Figura 5-10: Mapa de dades elementals, evidències i sistemes d'informació de l'objecte "Auditoria de protecció de dades personals"



Determinats elements, com el nivell de dadificació o la complexitat del model, es poden identificar ràpidament a través de la seva representació gràfica. Veiem doncs com en el primer cas el nivell de dadificació és molt alt, mentre que el segon depèn bàsicament de la generació d'evidències-objecte.

El mateix indicador el proposem també al nivell de superobjecte.

DAD-25		
Indicador de superobjecte	Índex de rellevància global	Interpretació de la mesura
Indicadors de base: DAD-15 Nombre de processos DAD-16 Nombre de sèries documentals DAD-14 Nombre d'evidències DAD-13 Nombre de sistemes d'informació	Fórmula de càlcul: $\frac{([DAD-15]+[DAD-16]+[DAD-14]+[DAD-13])Flt[ENT006-001]}{([DAD-15]+[DAD-16]+[DAD-14]+[DAD-13])/2}$	Un valor alt en aquest índex indica una major presència del superobjecte en un conjunt d'activitats, els documents que se'n generen i els sistemes d'informació que hi donen suport.

5.5.4 Mapa de relacions entre dades i objectes

Per a l'anàlisi visual del mapa de dades, es poden generar visualitzacions que mostrin de forma interactiva les relacions entre les dades, els objectes i els superobjectes. També es poden generar visualitzacions limitades a les relacions entre objectes, i desenvolupar-les de forma ontològica, en forma de *knowledge graph*. Per un exemple aplicat a les dades i els objectes vegeu el cas d'aplicació al Consorci d'Educació de Barcelona a l'apartat 6.7.4 del capítol 6, i per un exemple d'aplicació global vegeu l'apartat 6.7.5.

5.6 Quadre-resum dels indicadors

Taula 5-20: Taula resum de tots els indicadors del sistema de governança de la informació

	SÈRIE	PROCÉS	SISTEMA	DADA			
	Indicadors de sèrie documental	Indicadors de procés	Indicadors de sistema d'informació	Indicadors de dada elemental	Indicadors d'objecte	Indicadors de super objecte	Indicadors de tipus d'objectes
Nombre de sèries documentals	SER-01	PRO-04	SIS-06	DAD-08	DAD-12	DAD-16	DAD-20
Nombre d'evidències	SER-02	PRO-03	SIS-07	DAD-06	DAD-10	DAD-14	DAD-18
Nombre de regles de conservació i eliminació	SER-03						
Nombre de regles d'accés i ús	SER-04						
Nombre de sèries documentals relacionades	SER-05						
Nombre de dades elementals	SER-06	PRO-07	SIS-09	DAD-01			
Nombre d'ocurrències de dades	SER-07	PRO-08	SIS-10	DAD-02			
Nombre d'objectes	SER-08	PRO-09	SIS-11	DAD-03			
Nombre de superobjectes	SER-09	PRO-10	SIS-12	DAD-04			
Índex d'influència simple / de criticitat (sistemes d'informació)	SER-10	PRO-11	SIS-13				
Índex d'influència ponderat	SER-11	PRO-12					
Índex d'influència net	SER-12	PRO-13					
Índex de dadificació simple a nivell d'evidència	SER-13	PRO-14					
Índex de dadificació simple a nivell de dada elemental	SER-14	PRO-15					
Índex de dadificació ponderat	SER-15	PRO-16					
Índex de ròssec intraentitat	SER-16	PRO-17					
Índex de ròssec interentitat	SER-17	PRO-18					
Índex de transaccionalitat a partir de les dades elementals	SER-18		SIS-17				
Índex de transaccionalitat a partir del tipus de dades	SER-19		SIS-18				
Índex de transaccionalitat a partir dels objectes	SER-20		SIS-19				
Nombre de processos		PRO-01	SIS-05	DAD-07	DAD-11	DAD-15	DAD-19
Nombre de tasques		PRO-02					

	SÈRIE	PROCÉS	SISTEMA	DADA			
	Indicadors de sèrie documental	Indicadors de procés	Indicadors de sistema d'informació	Indicadors de dada elemental	Indicadors d'objecte	Indicadors de super objecte	Indicadors de tipus d'objectes
Nombre de sistemes d'informació		PRO-05	SIS-01	DAD-05	DAD-09	DAD-13	DAD-17
Nombre de processos relacionats		PRO-06					
Nombre de sistemes d'informació vigents			SIS-02				
Nombre de sistemes d'informació desenvolupats			SIS-03				
Nombre de sistemes d'informació deprecats			SIS-04				
Nombre de relacions amb sistemes d'informació			SIS-08				
Índex d'influència			SIS-14				
Índex de connectivitat ponderada			SIS-15				
Índex de connectivitat potencial			SIS-16				
Classificació per nivell de seguretat de les dades			SIS-20				
Índex de rellevància dins el procés o la sèrie documental				DAD-21			
Índex de rellevància dins el sistema d'informació				DAD-22			
Índex de rellevància global				DAD-23	DAD-24	DAD-25	

Capítol 6. Cas d'aplicació: el sistema de governança de la informació del Consorci d'Educació de Barcelona

6.1 L'objecte d'anàlisi: el Consorci d'Educació de Barcelona

Com s'ha descrit en l'apartat 1.5.2.3 de la introducció (capítol 1), El Consorci d'Educació de Barcelona és l'ens del sector públic català que s'ocupa de la gestió, ordenació i millora de l'educació no universitària a la ciutat de Barcelona. El seu abast competencial inclou tots els centres educatius de la ciutat de Barcelona (públics i privats, amb concert o sense), i tots els nivells educatius no universitaris.

El Departament d'Educació de la Generalitat de Catalunya és l'òrgan competent per a la gestió, ordenació, inspecció i millora de l'educació no universitària a tot el territori català, tal com estableix la Llei 12/2009, del 10 de juliol, d'educació. Desplega les seves competències de forma directa, a través de la pròpia estructura departamental, de la Inspecció d'educació (Capítol VII de la Llei 12/2009), dels serveis territorials distribuïts per cada àrea territorial (Article 175 de la Llei 12/2009), i d'una sèrie d'ens del sector públic, com el Consell Superior de l'Avaluació (creat pel Decret 305/1993, de 9 de desembre, de creació del Consell Superior d'Avaluació del Sistema Educatiu).

El Departament d'Educació conforma, juntament amb els ens locals, el que s'anomena Administració educativa (Article 156 de la Llei 12/2009).²⁴³ La ciutat de Barcelona gaudeix d'un règim especial, reconegut per l'Article 160 de la Llei 12/2009, i regulat per la Llei 22/1998, del 30 de desembre, de la Carta municipal de Barcelona. El Consorci d'Educació de Barcelona actua doncs en qualitat d'Administració educativa a la ciutat de Barcelona, i inclou al mateix temps les funcions de servei territorial de Barcelona ciutat. Únicament les escoles bressol municipals queden fora d'aquest abast, atès que són gestionades per l'Institut Municipal d'Educació de Barcelona (IMEB), creat per l'Ajuntament de Barcelona l'any 1993 com un organisme autònom local.

²⁴³ Article 156 Administració educativa.

1. L'Administració educativa és l'Administració de la Generalitat i actua per mitjà del Departament.
2. Els ens locals tenen la condició d'Administració educativa en l'exercici de les competències pròpies, d'acord amb l'Estatut, i exerceixen també les competències que els són atribuïdes d'acord amb el que estableix aquesta llei.

Els serveis territorials del Departament d'educació despleguen la seva activitat sobre el territori tant de forma directa com a través dels serveis educatius de zona (Article 86 de la Llei 12/2009). Dins els serveis educatius s'hi troben els Equips d'assessorament i orientació psicopedagògica (EAP), els Centres de recursos pedagògics (CRP), i diferents perfils professionals de suport als centres educatius (treballadors/ores socials, tècnics d'integració social per a l'atenció a la diversitat, assessors de llengua i cohesió social, etc.). En el cas del Consorci d'Educació de Barcelona, disposa d'un servei educatiu a cada districte de la ciutat, formant una xarxa on treballen aproximadament 130 professionals.

Els centres educatius, tant els públics com els privats, son autònoms pel que fa a la seva organització interna i desenvolupament de les seves funcions, tal estableix el Decret 102/2010, de 3 d'agost, d'autonomia dels centres educatius.

Als efectes de lectura funcional, doncs, determinem l'abast de la recollida de dades a partir del marc exposat:

- El Departament d'educació, juntament amb la Inspecció d'educació, que en depèn directament, i els ens del sector públic que hi estan vinculats, exerceixen un conjunt de competències exclusives que generen processos propis.
- L'acció sobre el territori es desenvolupa a través dels serveis territorials, i dels serveis educatius que depenen de cada servei territorial. Desenvolupen per tant una sèrie de processos propis, però governats del Departament.
- Els centres educatius, com a resultat del seu marc d'autogestió, desenvolupen també una sèrie de processos propis.

El Consorci d'Educació de Barcelona s'ubica doncs en el segon epígraf, atès que, en tant que servei territorial, desenvolupa un conjunt de processos propis de governança departamental. Però pel seu caràcter singular, hi afegeix alguns processos més que es deriven de les competències que li atribueix la Carta municipal de Barcelona, i pel mateix motiu gaudeix d'un cert marge de llibertat en la definició dels processos governats pel Departament.

L'abast de la recollida i anàlisi de dades s'ha centrat doncs en els processos propis del Consorci d'Educació de Barcelona i dels deu serveis educatius de districte de Barcelona ciutat. Els processos del Departament, de la Inspecció d'educació, i dels centres educatius, dels quals també s'ha recollit informació, queden fora de l'abast d'aquesta anàlisi.

6.2 Aplicació de la metodologia i els indicadors al Consorci d'Educació de Barcelona

6.2.1 Objectius del cas d'aplicació

Amb l'objectiu de posar a prova la metodologia de mapatge i dadificació (**GMAP** i **GDAD**), el model de dades i els indicadors d'anàlisi proposats en els capítols 3, 4 i 5, s'ha implantat al Consorci d'Educació de Barcelona un sistema de governança de la informació basat en aquest model.

Els objectius d'aquesta implantació han estat els següents:

- Valorar la factibilitat d'implantar el model, obtenir la informació necessària per alimentar-lo i mantenir-la actualitzada.
- Avaluar la comprensibilitat de la metodologia per part de professionals de la gestió documental i de l'anàlisi de processos.
- Dimensionar l'esforç necessari per aplicar la metodologia, recollir les dades i mantenir-les actualitzades, i valorar els punts que poden incidir en l'increment o decrement d'aquest esforç.
- Obtenir informació quantitativa i qualitativa per a la verificació o refutació de les hipòtesis plantejades en aquesta tesi.

La valoració d'aquests objectius s'ha inclòs al capítol 8 d'aquesta tesi.

6.2.2 Base tecnològica del cas d'aplicació

Com s'ha descrit en el capítol 4, el sistema de governança de la informació del Consorci d'Educació de Barcelona utilitza com a base tecnològica l'aplicació informàtica CATPROC 2.0, en la forma i amb les funcionalitats ja relacionades.

Els indicadors de governança s'han implementat físicament en quatre quadres de comandament, un per cada entitat analitzada. Cada quadre disposa d'una o varies pàgines per a cada indicador, i permet combinar nivells d'agregació i filtres per obtenir mesures agregades o parcials. Els quadres també incorporen visualitzacions interactives dels mapes de relacions entre entitats i de les matrius de variables, com també, en alguns casos, correlacions de variables o anàlisi interactiva d'influenciadors clau.

Els quadres de comandament estan desenvolupats en *Microsoft Power BI*,²⁴⁴ i obtenen les seves dades de l'aplicació CATPROC 2.0, a la qual estan connectats directament. Amb cada actualització les dades es depuren, es transformen i es carreguen al model de dades del quadre de comandament a través de processos ETL (*Extract, Transform, Load*) desenvolupats en *Microsoft Power Query*.²⁴⁵

Per al càlcul dels indicadors s'ha utilitzat les funcions de relació, filtrat i agregació de *Power BI*, juntament amb sentències en llenguatge DAX,²⁴⁶ i en alguns casos càlculs en R integrats en visuals de *Power BI*.²⁴⁷ Per a les visualitzacions s'ha completat la paleta estàndard de *Power BI* amb visuals interactius de *ZoomCharts*.²⁴⁸

Alguns dels càlculs realitzats per a les conclusions d'aquest projecte han requerit utilitzar també taules dinàmiques de *Microsoft Excel*, i per a les anàlisi de *clustering* i correlació de variables en alguns casos s'ha utilitzat el programari *Orange Data mining*, en la seva versió 3.35.²⁴⁹

6.3 Estat de la recollida de dades

La fase de recollida de dades del projecte s'ha desenvolupat des del febrer de 2017, en què va començar el mapatge, fins el maig de 2023, en què s'ha tancat la vista dels mapes sobre la qual s'han elaborat els indicadors d'anàlisi. En aquest temps, el Consorci ha estat objecte de dues reestructuracions orgàniques parcials, i de canvis constants en els seus processos de treball i infraestructura tecnològica. Els principals canvis que han afectat el Consorci durant aquest període han estat els següents:

²⁴⁴ Vegeu: <https://powerbi.microsoft.com/es-es/>. Les versions de Microsoft Power BI utilitzades s'han actualitzat diverses vegades al llarg del projecte.

²⁴⁵ Vegeu: <https://powerquery.microsoft.com/en-us/>.

²⁴⁶ DAX és un llenguatge de expressions de fórmules propi de *Microsoft SQL Server Analysis Services* (SSAS), i també utilitzat a *Microsoft Power Pivot* per a *Excel*.

²⁴⁷ La versió de R implementada en com a component en *Power BI* és la versió *Microsoft R 3.4.4*.

²⁴⁸ Vegeu: <https://zoomcharts.com/en/microsoft-power-bi-custom-visuals/>.

²⁴⁹ *Orange Data mining* és una eina *Open source* de visualització i anàlisi de dades, disponible a <https://orangedatamining.com/>.

Taula 6-1: Canvis que han afectat al Consorci d'Educació de Barcelona durant el projecte

Canvi	Data del canvi
Posada en marxa del Pla de Xoc contra la Segregació Escolar i modificació de la distribució i reserva de places durant el procés de preinscripció.	Abril de 2017
Canvi en l'adscripció de competències sobre l'Escolarització a la Direcció d'Atenció a la Comunitat Educativa.	Maig de 2018
Modificació de les bases del procés de gestió d'ajuts menjador i vinculació a dades de l'Institut Municipal de Serveis Socials.	Juliol de 2018
Inici de la transformació dels serveis informàtics del CEB cap al model del Centre de Telecomunicacions i Tecnologies de la Informació de la Generalitat de Catalunya.	Maig de 2019
Creació de la Direcció d'Ensenyaments Post-obligatoris i transferència de les competències globals sobre l'educació en l'etapa post-obligatòria.	Setembre de 2019
Publicació del Decret 11/2021, de 16 de febrer, de la programació de l'oferta educativa i del procediment d'admissió en els centres del Servei d'Educació de Catalunya.	Novembre de 2021
Reestructuració dels àmbits clau del Consorci i creació de la Direcció de Territori, i modificació de competències i processos de l'Àrea d'Educació Inclusiva.	Gener de 2022

A part dels canvis externs, el propi procés de recollida ha vist modificat el seu abast, en funció dels resultats que s'anaven obtenint i del potencial que s'observava en aquestes dades. Els canvis d'abast i plantejament han estat els següents:

Taula 6-2: Canvis d'abast i plantejament en la recollida de dades del projecte

Canvi	Data del canvi
Inclusió dels processos i sèries dels serveis educatius de districte.	Abril de 2017
Redefinició de la granularitat en la identificació de sistemes d'informació amb el desplegament de l'entorn Microsoft 365 amb Power Platform.	Abril de 2020
Exploració de la inclusió dels processos i sèries dels centres educatius.	Juliol de 2020
Creació de l'entitat ENT013 Visualització de dades i inventari sistemàtic dels quadres de comandament.	Juliol de 2019
Segregació del concepte d'objecte descrit en la descripció de dades elementals, i creació de l'entitat ENT014 Objecte.	Desembre de 2020

En el moment del tancament de la vista dels mapes sobre la qual s'han calculat els valors de els indicadors d'anàlisi, la situació quantitativa per cadascuna de les entitats, considerant exclusivament l'abast Consorci d'Educació de Barcelona, és la següent:

Taula 6-3: Total d'instàncies per entitat a maig de 2023

	Data de primera alta a CATPROC	Total instàncies identificades	Instàncies vigents a maig de 2023
ENT001 Agent	25/02/2017	116	115
ENT002 Procés	24/02/2017	250	217
ENT003 Tasca ²⁵⁰	10/12/2018	688	688
ENT004 Evidència	11/01/2017	1.332	1.239
ENT005 Model de document	08/05/2017	694	412
ENT006 Dada elemental	12/06/2019	2.291	2.291
ENT007 Sèrie documental	15/03/2017	498	460
ENT008 Regla d'accés	15/03/2017	1.053	1.029
ENT009 Regla de conservació	15/03/2017	362	357
ENT010 Disposició normativa	23/09/2019	26	26
ENT011 Sistema d'informació	24/02/2017	658	451
ENT012 Interfície de connexió	28/08/2019	277	263
ENT013 Visualització de dades	10/07/2019	199	103
ENT014 Objecte	04/12/2020	171	171

Els canvis viscuts pel Consorci durant l'execució del projecte, i els canvis derivats de l'evolució del propi projecte, han fet molt difícil establir parts dels diferents mapes d'entitats. En alguns casos ha estat possible aconseguir un mapatge exhaustiu, però en d'altres les condicions de context no ho ha permès. De les entitats que s'identifiquen amb l'enfocament *top-down* (**GMAP**), el percentatge de dades recollides és el següent:

Taula 6-4: Percentatge de dades recollides per entitat

	Percentatge d'instàncies identificades i caracteritzades	Relacions entre les instàncies de cada entitat
ENT001 Agent	100%	100%
ENT002 Procés	100%	100%
ENT005 Model de document	100%	No aplica
ENT007 Sèrie documental	80%	100%
ENT008 Regla d'accés	65%	No aplica
ENT009 Regla de conservació	61%	No aplica
ENT010 Disposició normativa	25% ²⁵¹	No aplica

²⁵⁰ El buidatge de les tasques no s'ha pogut realitzar sistemàticament. Les tasques s'han informat preferentment en aquells processos que han estat objecte de redisseny i transformació.

²⁵¹ Sobre el total estimat.

	Percentatge d'instàncies identificades i caracteritzades	Relacions entre les instàncies de cada entitat
ENT011 Sistema d'informació	100%	50% ²⁵²
ENT013 Visualització de dades	100%	No aplica

L'enfocament *bottom-up* s'ha portat a terme en totes les actuacions de millora i dadificació de processos (**GDAD**), que al llarg d'aquest període de recollida de dades han afectat total o parcialment processos de totes les famílies.

Taula 6-5: Percentatge de processos dadificats per família de processos

	Identificació i caracterització de tasques (ENT003)	Identificació i caracterització d'evidències (ENT004)	Identificació i caracterització de dades i objectes (ENT006/ENT014)
Assegurament compliment RGPD	Parcial	100%	100%
Atenció a les obligacions de transparència	Parcial	100%	100%
Gestió documental i arxiu	100%	75%	63%
Racionalització i millora organitzativa	Parcial	100%	100%
Anàlisi i producció d'informació	100%	89%	67%
Atenció a les necessitats educatives especials	Parcial	79%	21%
Escolarització	Parcial	100%	50%
Resta de famílies de processos	Parcial	Parcial	Parcial

El fet de no poder assolir l'exhaustivitat en la recollida de dades de totes les entitats en un model on totes les entitats estan relacionades entre elles ens ha obligat, en la fase d'anàlisi, a combinar lectures globals d'entitats en concret, amb lectures parcials de determinats conjunts de relacions. Així doncs, als efectes de poder fer una anàlisi detallada de les característiques dels processos dadificats es requereix disposar de l'anàlisi completa, incloent els sistemes d'informació i l'assignació completa de dades tant a evidències com a sistemes, i d'una estabilitat en el temps del procés dadificat. Per aquest motiu, s'ha triat un **grup focal** amb els processos que compleixen aquestes dues

²⁵² Sobre el total estimat d'interaccions. En el cas de les interaccions entre sistemes d'informació, materialitzades a través de l'entitat **ENT012 Interfície de connexió**, s'ha recollit la informació del 100% d'interaccions amb el *data lake* del Consorci d'Educació de Barcelona (projecte "Banc de Dades"), i del 100% d'interaccions entre els sistemes vinculats a l'escolarització en general (família de processos "Escolarització"), i a l'escolarització de l'alumnat amb necessitats específiques de suport educatiu (família de processos "Atenció a les necessitats educatives especials"). En la resta de casos ha estat impossible obtenir informació tècnica suficient per a descriure amb precisió les interaccions.

característiques, que ha estat objecte d'un estudi més detallat. Els 27 processos que s'inclouen en el grup focal són els següents:

Taula 6-6: Processos inclosos en el grup focal

PROCESSOS	EVIDÈNCIES	OCURRÈNCIES
DOC001 Desenvolupament del sistema de gestió documental	11	62
DOC002 Tractament documental	8	46
DOC003 Transferències a l'Arxiu Central	9	74
DOC004 Atenció a préstecs i consultes de documentació	16	83
ESC001 Seguiment de falsos empadronaments	10	112
GRH012 Gestió de les vagues (relacions laborals)	3	20
GRH023 Gestió de borses de personal de suport als centres educatius (COVID-19)	6	58
INF001 Gestió, manteniment i ús del banc de dades del CEB	17	183
INF002 Comunicació de les eleccions dels Consells Escolars	1	36
INF007 Coordinació de l'ús social de centres	2	34
INF009 Manteniment de les dades de la Base de Dades de Centres (BDC)	6	134
LOP001 Desenvolupament del sistema de protecció de dades de caràcter personal	5	26
LOP002 Atenció al drets d'autodeterminació informativa reconeguts per la normativa de protecció de dades personals	3	30
LOP003 Gestió de les incidències de seguretat en protecció de dades	6	64
LOP005 Avaluació i control de la protecció de dades (auditories)	9	67
LOP006 Manteniment i publicació del registre d'activitats de tractament	3	19
LOP007 Assessorament en matèria de protecció de dades personals	5	25
NEE002 Actuacions en casos de sospita de maltractament i/o assetjament en un centre educatiu	6	134
NEE020 Intervencions en matèria de seguretat ciutadana i prevenció de radicalismes violents	1	15
NEE021 Reconeixement d'alumnat amb NEE-B	8	88
ORG002 Gestió de programes en l'àmbit de processos i sistemes	4	15
ORG003 Anàlisi i millora de processos	9	38
ORG004 Normalització de formularis i models de documents	14	135
TRA001 Publicació d'informació a l'espai de transparència	2	5

PROCESSOS	EVIDÈNCIES	OCURRÈNCIES
TRA002 Atenció a les peticions d'accés a documents i dades del CEB de la part interessada	13	85
TRA003 Assessorament en matèria de transparència als centres educatius	2	16
TRA004 Atenció a les peticions d'accés a documents i dades del CEB	13	87

No totes les sèries documentals derivades dels processos del grup focal han pogut ser objecte de dadificació. Per aquest motiu, només s'ha inclòs al grup focal les 32 sèries documentals que sí que compleixen el requisit de recollir la part realment dadificada de cada procés. Les sèries documentals del grup focal són les següents:

Taula 6-7: Sèries documentals incloses en el grup focal

SÈRIES DOCUMENTALS	EVIDÈNCIES	OCURRÈNCIES
DOC001-001 Desenvolupament del sistema de gestió documental	6	22
DOC001-002 Assessorament als centres i serveis educatius en matèria de gestió documental	5	40
DOC002-001 Tractament documental	8	46
DOC003-001 Transferències a l'Arxiu Central	9	74
DOC004-001 Préstecs i consultes de documentació	16	83
ESC001-001 Seguiment de falsos empadronaments	10	112
GRH012-004 Comunicació d'exercici del dret de vaga	3	20
GRH023-001 Gestió de borses de personal de suport als centres educatius (COVID-19)	6	58
INF001-001 Explotació de dades	8	111
INF001-002 Manteniment del banc de dades del CEB	5	43
INF001-003 Elaboració d'estudis amb col·laboracions externes	4	29
INF002-001 Eleccions al Consell Escolar de Centre	1	36
INF007-001 Coordinació de l'ús social de centres	2	34
INF009-001 Manteniment de les dades de la Base de Dades de Centres (BDC)	5	129
INF009-002 Manteniment de les dades de la Base de dades de Centres (BDC)	1	5
LOP001-001 Desenvolupament del sistema de protecció de dades de caràcter personal	5	26
LOP002-001 Atenció al drets d'autodeterminació informativa reconeguts per la normativa de protecció de dades personals	3	30
LOP003-001 Registre i notificacions de violacions de seguretat	6	64

SÈRIES DOCUMENTALS	EVIDÈNCIES	OCURRÈNCIES
LOP005-001 Auditories APDCAT i controls de seguretat sobre protecció de dades de caràcter personal	9	67
LOP006-001 Manteniment i publicació del registre d'activitats de tractament	3	19
LOP007-001 Assessorament en matèria de protecció de dades personals	5	25
NEE002-001 Actuacions en casos de sospita de maltractament i/o assetjament en un centre educatiu	6	134
NEE020-001 Intervencions en matèria de seguretat ciutadana i prevenció de radicalismes violents	1	15
NEE021-001 Seguiment i valoració d'alumnes en situació de vulnerabilitat i/o malestar social o psicosocial	8	88
ORG002-001 Gestió de programes en l'àmbit de processos i sistemes	4	15
ORG003-001 Anàlisi i millora de processos	9	38
ORG004-001 Normalització de formularis i models de documents	12	124
ORG004-002 Normalització de formularis i models de document	2	11
TRA001-001 Publicació d'informació a l'espai de transparència	2	5
TRA002-001 Gestió de les sol·licituds d'accés a documents i dades per part interessada	13	85
TRA003-001 Assessorament en matèria de transparència als centres educatius	2	16
TRA004-001 Atenció a les peticions d'accés a la informació pública	13	87

El mapa de dades disponible en la vista de tancament dels mapes s'alimenta d'una segona font, les dades assignades directament a sistemes d'informació. Cal tenir present que la identificació d'objectes es fa a partir de les dades elementals detectades i assignades a sistemes d'informació o evidències. Les mètriques d'ocurrències de dades en sistemes i evidències són les següents:

Taula 6-8: Nombre d'ocurrències de dades en sistemes i evidències

	Sistemes d'informació (ENT011)	Evidències (ENT004)
Ocurrències de dades elementals assignades al total d'entitats mapades	3.334	2.013
Ocurrències de dades assignades a les entitats vigents de cada àmbit	2.779	2.008

La informació que s'ha pogut recollir ha condicionat, tant per l'exhaustivitat com per la fiabilitat, els indicadors que es podien aplicar. Per aquest motiu, en el cas d'aplicació del Consorci no s'han calculat tots els indicadors del model, sinó aquells que podien proporcionar una informació més fiable i significativa, fent especial èmfasi en aquells vinculats directament a la confirmació o desestimació de les hipòtesis plantejades en els objectius de la tesi.

Passem a continuació a descriure el resultat de l'aplicació de cadascun dels indicadors que proposem per al model de governança en el cas del Consorci d'Educació de Barcelona. L'anàlisi ha estat més exhaustiu en aquells casos on l'indicador s'ha utilitzat per a la validació d'hipòtesis a les conclusions (sèries documentals i sistemes d'informació), mentre que en la resta de casos (processos i dades elementals) s'ha limitat bàsicament a les mesures descriptives.

6.4 Indicadors sobre les sèries documentals

6.4.1 Mètriques descriptives de les sèries documentals

SER-01 Nombre de sèries documentals

A dia d'avui el Consorci d'Educació de Barcelona té identificades 460 sèries documentals vigents, que es distribueixen per tipologia i productor de la forma següent:

Taula 6-9: Sèries documentals per productor i tipologia de sèrie

CARÀCTER	PRODUCTOR: CONSORCI (CEB)	PRODUCTOR: CENTRES EDUCATIUS	TOTAL DE SÈRIES DOCUMENTALS
Primàries	338	79	417
Secundàries	41	2	43
Total	379	81	460

En aquesta anàlisi ens centrarem en les sèries documentals produïdes directament pel Consorci d'Educació de Barcelona, que inclouen les que produeixen les unitats administratives de serveis centrals, i les dels serveis educatius de districte. En total seran objecte d'anàlisi 379 sèries documentals.

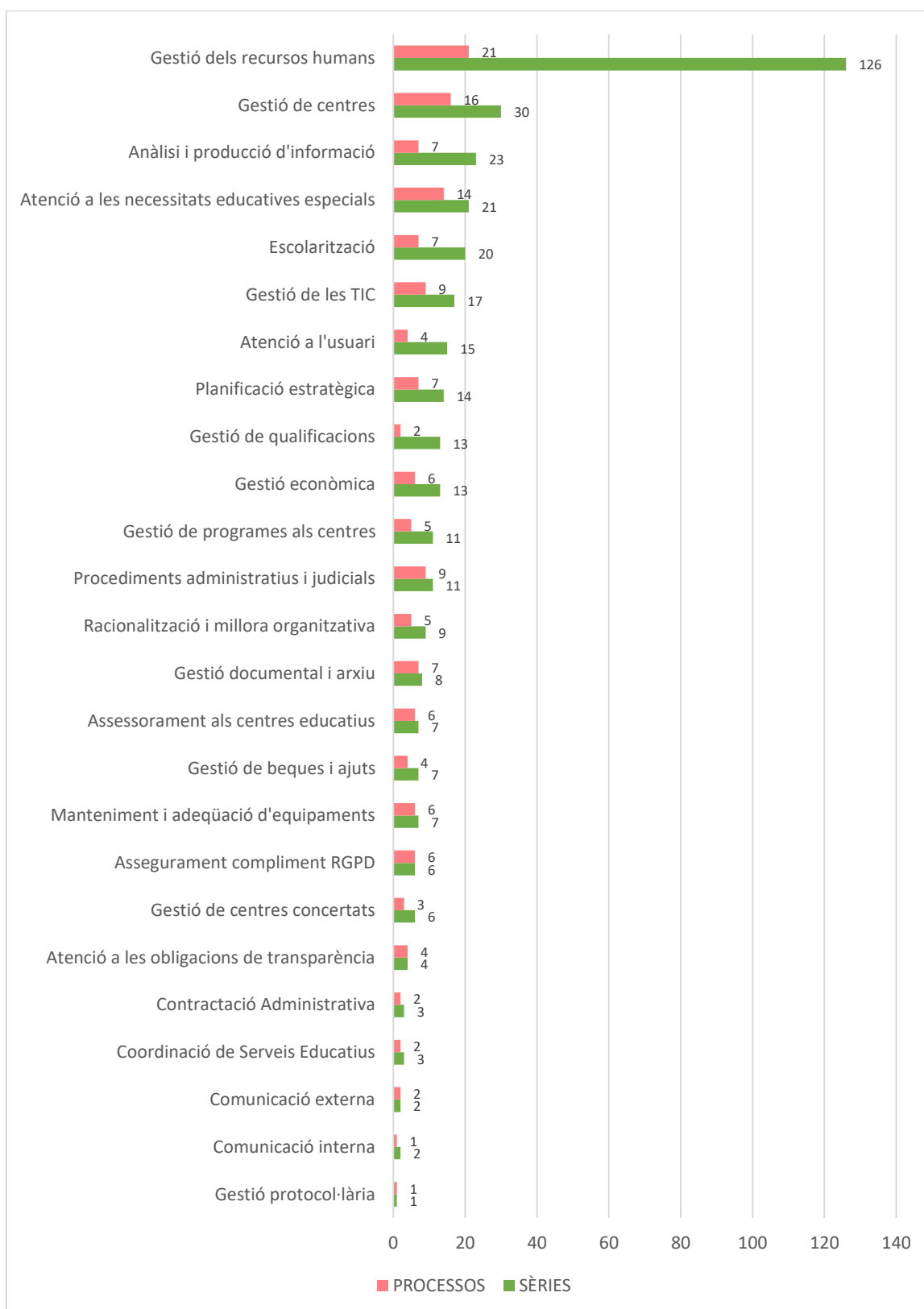
Si analitzem la distribució d'aquestes sèries en funció del tipus de procés generador, el resultat és el següent:

Taula 6-10: sèries documentals per tipus de procés

TIPUS DE PROCÉS	SÈRIES DOCUMENTALS
Estratègic	25
Clau	137
De Suport	217

I si fem la mateixa distribució en funció de les famílies de processos, la distribució és la següent:

Gràfic 6-1: Nombre de processos i sèries documentals per família de processos



Si calculem la proporció entre processos i sèries per famílies de processos observem el següent resultat:

Taula 6-11: Nombre de sèries per procés i família de processos

FAMÍLIA	SÈRIES PER PROCÉS
Gestió dels recursos humans	6,0
Gestió de centres	1,9
Anàlisi i producció d'informació	3,3
Atenció a les necessitats educatives especials	1,5
Escolarització	2,9
Gestió de les TIC	1,9
Atenció a l'usuari	3,8
Planificació estratègica	2,0
Gestió econòmica	2,2
Gestió de qualificacions	6,5
Procediments administratius i judicials	1,2
Gestió de programes als centres	2,2
Racionalització i millora organitzativa	1,8
Gestió documental i arxiu	1,1
Manteniment i adequació d'equipaments	1,2
Gestió de beques i ajuts	1,8
Assessorament als centres educatius	1,2
Gestió de centres concertats	2,0
Assegurament compliment RGPD	1,0
Atenció a les obligacions de transparència	1,0
Coordinació de Serveis Educatius	1,5
Contractació Administrativa	1,5
Comunicació interna	2,0
Comunicació externa	1,0
Gestió protocol·lària	1,0

Del total dels processos del CEB, s'identifiquen 61 processos que no tenen encara cap sèrie documental identificada. Per tant, les sèries que s'han identificat corresponen a un total de 156 processos. Observem en aquest cas que la mitjana de sèries per procés és de 2,1, i si eliminem els dos *outliers*²⁵³ (Gestió dels recursos humans i Gestió de qualificacions), la mitjana es redueix tres punts, fins al 1,8. En les conclusions fem una anàlisi més detallada d'aquest indicador per tal de determinar la rellevància real de la diferència entre els conceptes de procés i de sèrie documental.

²⁵³ Gestió dels recursos humans i Gestió de qualificacions generen sèries per cadascuna de les modalitats de tràmit, en la qual el procediment és pràcticament el mateix però els models de document són diferents. Per aquest motiu generen una proporció major de sèries que pot distorsionar la mitjana general.

Per a l'anàlisi detallada de les característiques de les sèries documentals dadificades ens centrarem en les sèries 32 del grup focal. Aquestes sèries es caracteritzen per les següents mètriques:

Taula 6-12: Mètriques de referència de les sèries documentals del grup focal

SÈRIES DOCUMENTALS	EVIDÈNCIES	OCURRÈNCIES	DADES	OBJECTES
Anàlisi i producció d'informació				
INF001-001 Explotació de dades	8	111	44	6
INF001-002 Manteniment del banc de dades del CEB	5	43	43	10
INF001-003 Elaboració d'estudis amb col·laboracions externes	4	29	17	5
INF002-001 Eleccions al Consell Escolar de Centre	1	36	36	4
INF007-001 Coordinació de l'ús social de centres	2	34	22	5
INF009-001 Manteniment de les dades de la Base de Dades de Centres (BDC)	5	129	101	7
INF009-002 Manteniment de les dades de la Base de dades de Centres (BDC)	1	5	5	3
Assegurament compliment RGPD				
LOP001-001 Desenvolupament del sistema de protecció de dades de caràcter personal	5	26	22	4
LOP002-001 Atenció al drets d'autodeterminació informativa [...]	3	30	20	4
LOP003-001 Registre i notificacions de violacions de seguretat	6	64	40	6
LOP005-001 Auditories APDCAT [...]	9	67	27	4
LOP006-001 Manteniment i publicació del registre d'activitats de tractament	3	19	10	3
LOP007-001 Assessorament en matèria de protecció de dades personals	5	25	20	7
Atenció a les necessitats educatives especials				
NEE002-001 Actuacions en casos de sospita de maltractament i/o assetjament [...]	6	134	32	8
NEE020-001 Intervencions en matèria de seguretat ciutadana i prevenció [...]	1	15	15	3
NEE021-001 Seguiment i valoració d'alumnes en situació de vulner. [...]	8	88	48	9
Atenció a les obligacions de transparència				
TRA001-001 Publicació d'informació a l'espai de transparència	2	5	4	2
TRA002-001 Gestió de les sol·licituds d'accés a documents i dades per part interessada	13	85	26	6

SÈRIES DOCUMENTALS	EVIDÈNCIES	OCURRÈNCIES	DADES	OBJECTES
TRA003-001 Assessorament en matèria de transparència als centres educatius	2	16	8	5
TRA004-001 Atenció a les peticions d'accés a la informació pública	13	87	28	6
Escolarització				
ESC001-001 Seguiment de falsos empadronaments	10	112	36	11
Gestió dels recursos humans				
GRH012-004 Comunicació d'exercici del dret de vaga	3	20	17	7
GRH023-001 Gestió de borses de personal de suport als centres educatius (COVID-19)	6	58	25	4
Gestió documental i arxiu				
DOC001-001 Desenvolupament del sistema de gestió documental	6	22	8	3
DOC001-002 Assessorament als centres i serveis educatius en matèria de [...]	5	40	13	4
DOC002-001 Tractament documental	8	46	15	10
DOC003-001 Transferències a l'Arxiu Central	9	74	22	10
DOC004-001 Préstecs i consultes de documentació	16	83	33	9
Racionalització i millora organitzativa				
ORG002-001 Gestió de programes en l'àmbit de processos i sistemes	4	15	8	3
ORG003-001 Anàlisi i millora de processos	9	38	16	4
ORG004-001 Normalització de formularis i models de documents	12	124	30	4
ORG004-002 Normalització de formularis i models de document	2	11	11	4

6.4.2 Mètriques derivades de la relació de la sèrie documental amb altres entitats

SER-02 Nombre d'evidències

Del total de sèries documentals que actualment tenen informada més d'una evidència al mapa de sèries (111 sèries), podem obtenir unes mitjanes orientatives.²⁵⁴ El nombre mitjà d'evidències per sèrie és de 5,4 evidències. D'aquestes, un 26% correspon a interfícies per interaccionar amb sistemes d'informació (tipus "Dades"), mentre que un 74% correspon a evidències-objecte (tipus "Document").

En el cas del grup focal la mitjana d'evidències per sèrie és de 5,9 evidències. En aquest cas, la proporció entre tipus d'evidències està molt més equilibrada, atès que es tracta de sèries que en la majoria dels casos han passat per un procés de dadificació. Les interfícies amb sistemes representen un 46% del total d'evidències, mentre que la mitjana d'evidències-objecte és d'un 54%.

SER-03 Nombre de regles de conservació i eliminació

De les 274 sèries documentals vigents per a les quals s'ha definit regles de conservació i eliminació, la distribució de regles per sèrie és la següent:

Taula 6-13: Nombre de regles de conservació i eliminació per sèrie documental

REGLES	SÈRIES
1 regla	217
2 regles	56
3 regles	1

Pel que fa a la resolució d'aquestes regles, en total 332, es distribueixen de la forma següent:

²⁵⁴ Una part important de les sèries documentals només estan vinculades a una única evidència de tipus dades, a través de la qual s'expressa el seu vincle amb un sistema d'informació. Aquestes sèries es descarten, atès que no s'ha pogut analitzar el seu circuit evidencial real.

Taula 6-14: Nombre de regles de conservació i eliminació per tipus de resolució

RESOLUCIÓ	REGLES	SÈRIES ²⁵⁵
Conservació	177	168
Eliminació	104	104
Pendent de definir ²⁵⁶	51	51

SER-04 Nombre de regles d'accés i ús

En el cas de les regles d'accés i ús, s'han definit 954 regles, que afecten un total de 291 sèries documentals, i que es distribueixen de la forma següent:

Taula 6-15: Nombre de regles d'accés i ús per sèrie documental

REGLES	SÈRIES
1 regla	1
2 regles	87
3 regles	61
4 regles	126
5 regles	7
6 regles	7
7 regles	1
8 regles	1

El major nombre de regles en relació amb les regles de conservació i eliminació és conseqüència de l'estructura de les regles d'accés i ús. Cada regla defineix uns privilegis d'accés i d'ús dels actius de la sèrie documental per a un agent en concret i durant un període de temps. A més agents participin d'un procés, més regles es definiran. La definició d'aquestes regles és altrament la forma com s'eliminen les sèries documentals secundàries en els processos d'adificats, substituint-les per l'accés als actius de la sèrie primària.

Pel que fa a les resolucions d'aquestes 954 regles, es distribueixen de la forma següent:

Taula 6-16: Nombre de regles d'accés i ús per tipus de resolució

RESOLUCIÓ	REGLES	SÈRIES
Esborrar	8	3
Crear	299	288
Modificar	1	1
Visualitzar	646	291

²⁵⁵ Una sèrie documental que estigui afectada per una resolució de conservació parcial o mostreig tindrà com a mínim dues regles assignades: una de conservació i una d'eliminació.

²⁵⁶ Es tracta de regles que estan en procés d'aprovació per part de la Comissió Nacional d'Accés, Avaluació i Tria Documental, o bé per part de la Gerència del Consorci en cas de les regles internes.

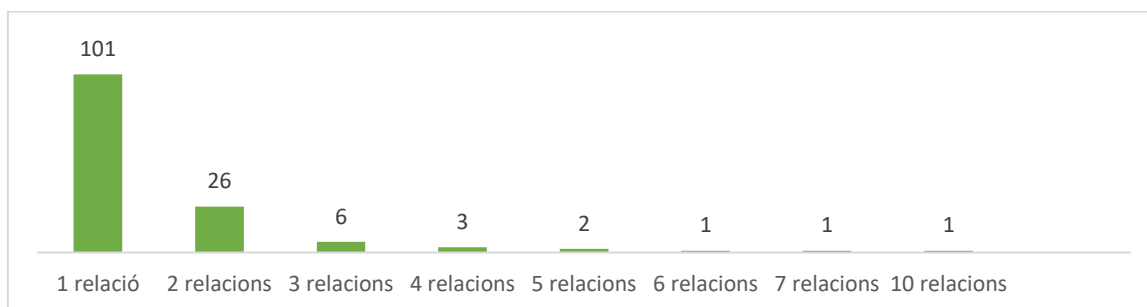
Cal tenir present que els permisos són incrementals, i es presenten en l'ordre de major a menor nombre de capacitats. Així, el permís d'esborrar permet també crear, modificar i visualitzar, el permís de crear també permet modificar i visualitzar, etc.

SER-05 Nombre de sèries documentals relacionades

La informació recollida durant el projecte només ha contemplat les relacions entre les sèries d'un mateix procés, especialment aquelles de continuïtat i dependència. No s'ha inclòs en aquesta recollida les relacions entre sèries de diferents processos, opció que s'incorpora amb l'indicador SER-12 (índex d'influència net) i s'explora amb l'indicador SER-17 (índex de ròssec intersèrie).

Un 31% de les 460 sèries documentals vigents té relacions actives amb altres sèries documentals del mateix procés, amb la següent distribució en el nombre de relacions per sèrie:

Gràfic 6-2: Nombre de sèries documentals pel nombre de relacions amb altres sèries



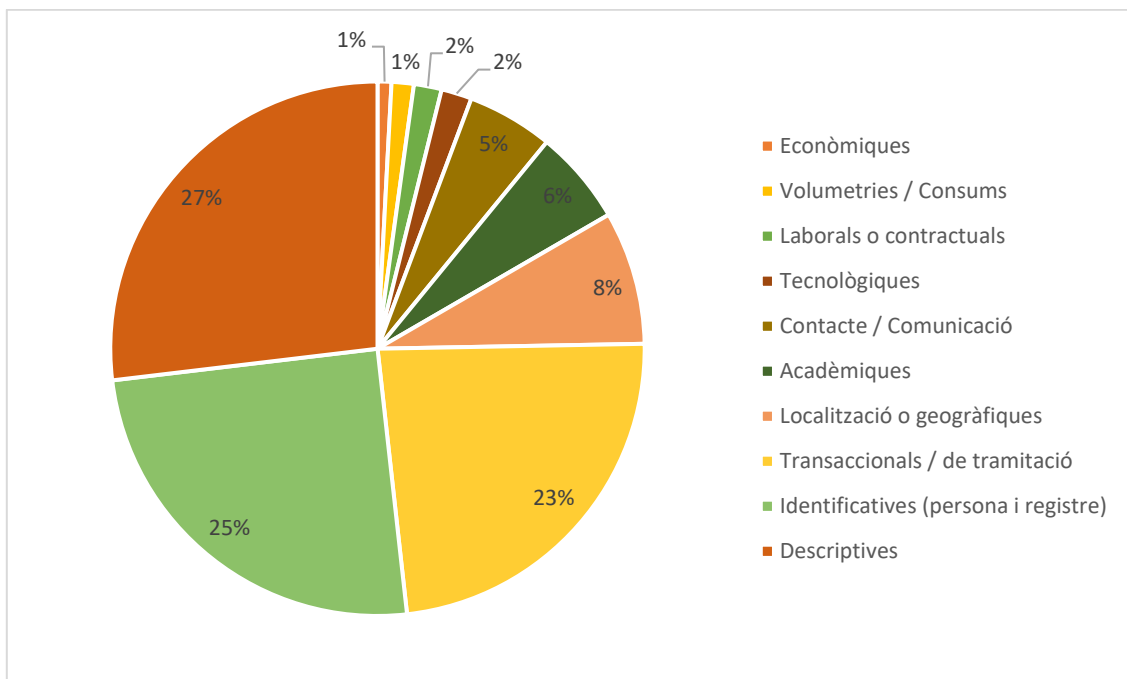
6.4.3 Mètriques sobre les dades de la sèrie documental

SER-06 Nombre de dades elementals

Per aquesta descripció ens centrarem en les sèries documentals del grup focal. El vincle entre la dada elemental i la sèrie documental s'obté del buidatge de les dades de les evidències vinculades a cada sèrie.

Dins les 32 sèries documentals del grup focal identifiquem 596 dades elementals diferents. La distribució d'aquestes dades en funció del tipus de dada oscil·la entre les següents magnituds:

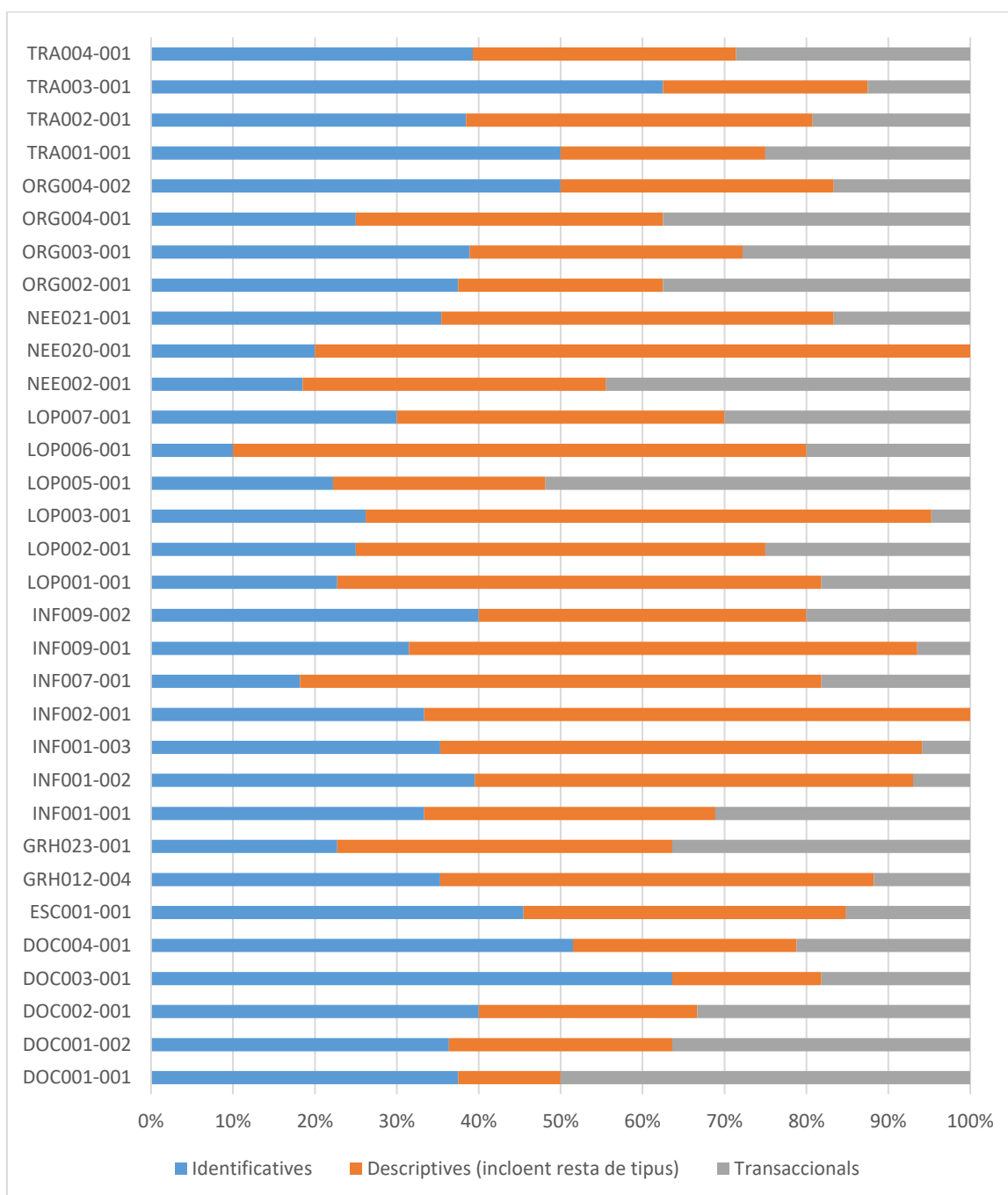
Gràfic 6-3: Percentatge de dades elementals per tipus de dada



Les categories que s'han fixat per a la caracterització de les dades obeeixen a usos i contextos d'utilització concrets de les dades. Podem considerar que les tres categories principals són les de dades identificatives, descriptives i transaccionals. La resta es poden encabir conceptualment dins la categoria de dades descriptives.

Si desglossem les dades de les sèries del grup focal per aquestes tres grans categories obtenim el següent:

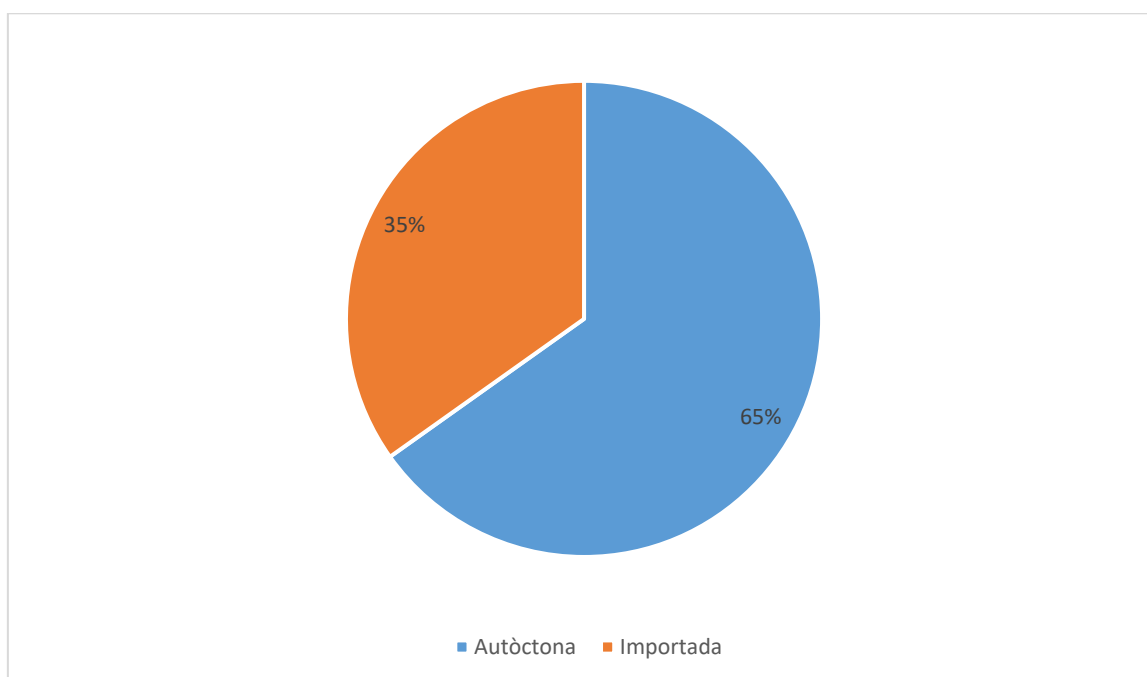
Gràfic 6-4: Percentatge de dades elementals per gran tipus i sèrie documental



SER-07 Nombre d'ocurrències de dades

Si ens centrem en les 32 sèries del grup focal, i incorporem el tipus d'ocurrència, observem que les 1691 ocurrències de dades es distribueixen de la forma següent:

Gràfic 6-5: Percentatge d'ocurrències de dades elementals per tipus d'ocurrència



Observem doncs que la proporció d'ocurrències és de 2 a 1 en favor de les ocurrències autòctones, cosa la qual denota una baixa integració de la informació en els processos dadificats, sobre un escenari teòric que probablement hauria de ser invers. Aquesta mesura ens indica que no existeix un nivell alt d'integració entre sistemes d'informació, i especialment que els sistemes d'informació no generen evidències de forma majoritàriament automatitzada, amb la qual cosa els usuaris finals es veuen obligats a reescriure en evidències-objecte dades de les quals ja disposen en sistema.

SER-08 Nombre d'objectes

Les mètriques per a les 32 sèries del grup focal són les següents:

Taula 6-17: Nombre d'objectes per tipus d'objecte i sèrie documental del grup focal

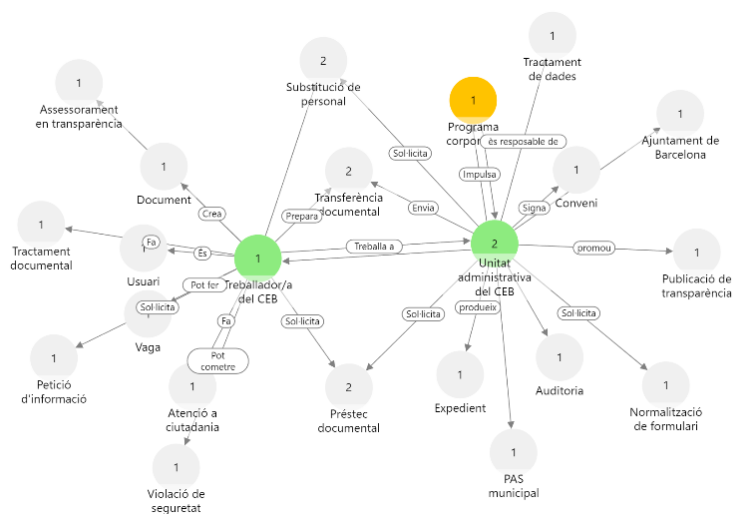
SÈRIE DOCUMENTAL	TIPUS ACCIÓ	TIPUS COSA
DOC001-001 Desenvolupament del sistema de gestió documental	1	2
DOC001-002 Assessorament als centres i serveis educatius en matèria de gestió docum.	1	3
DOC002-001 Tractament documental	4	6
DOC003-001 Transferències a l'Arxiu Central	3	7
DOC004-001 Préstecs i consultes de documentació	3	6
ESC001-001 Seguiment de falsos empadronaments	3	8
GRH012-004 Comunicació d'exercici del dret de vaga	3	4
GRH023-001 Gestió de borses de personal de suport als centres educatius (COVID-19)	1	3
INF001-001 Explotació de dades	1	5
INF001-002 Manteniment del banc de dades del CEB	4	6
INF001-003 Elaboració d'estudis amb col·laboracions externes	1	4
INF002-001 Eleccions al Consell Escolar de Centre	0	4

SÈRIE DOCUMENTAL	TIPUS ACCIÓ	TIPUS COSA
INF007-001 Coordinació de l'ús social de centres	2	3
INF009-001 Manteniment de les dades de la Base de Dades de Centres (BDC)	0	7
INF009-002 Manteniment de les dades de la Base de dades de Centres (BDC)	0	3
LOP001-001 Desenvolupament del sistema de protecció de dades de caràcter personal	2	2
LOP002-001 Atenció al drets d'autodeterminació informativa reconeguts per la [...]	1	3
LOP003-001 Registre i notificacions de violacions de seguretat	1	5
LOP005-001 Auditories APDCAT i controls de seguretat sobre protecció de dades de [...]	1	3
LOP006-001 Manteniment i publicació del registre d'activitats de tractament	2	1
LOP007-001 Assessorament en matèria de protecció de dades personals	3	4
NEE002-001 Actuacions en casos de sospita de maltractament i/o assetjament [...]	1	7
NEE020-001 Intervencions en matèria de seguretat ciutadana i prevenció de [...]	1	2
NEE021-001 Seguiment i valoració d'alumnes en situació de vulnerabilitat i/o [...]	4	5
ORG002-001 Gestió de programes en l'àmbit de processos i sistemes	1	2
ORG003-001 Anàlisi i millora de processos	2	2
ORG004-001 Normalització de formularis i models de documents	1	3
ORG004-002 Normalització de formularis i models de document	1	3
TRA001-001 Publicació d'informació a l'espai de transparència	1	1
TRA002-001 Gestió de les sol·licituds d'accés a documents i dades per part interessada	2	4
TRA003-001 Assessorament en matèria de transparència als centres educatius	1	4
TRA004-001 Atenció a les peticions d'accés a la informació pública	2	4

Observem que amb aquest indicador es compleix una constant: el nombre d'objectes de tipus "Cosa" és pràcticament sempre superior al nombre d'objectes de tipus "Acció". Aquesta és una regla esperable, atès que, com hem exposat en la **GDAD**, tot procés, i les sèries que se'n deriven, tenen un plantejament discursiu: un o diversos agents realitzen (preferentment) una acció que afecta un o diversos agents destinataris. La proporció invertida pot ser anòmla o relativa a casos singulars, mentre que l'absència d'accions, com veurem en les conclusions, podria ser indicador de sèries no transaccionals. Aquesta casuística s'analitza en l'objectiu 5 (apartat 7.6 del capítol 6).

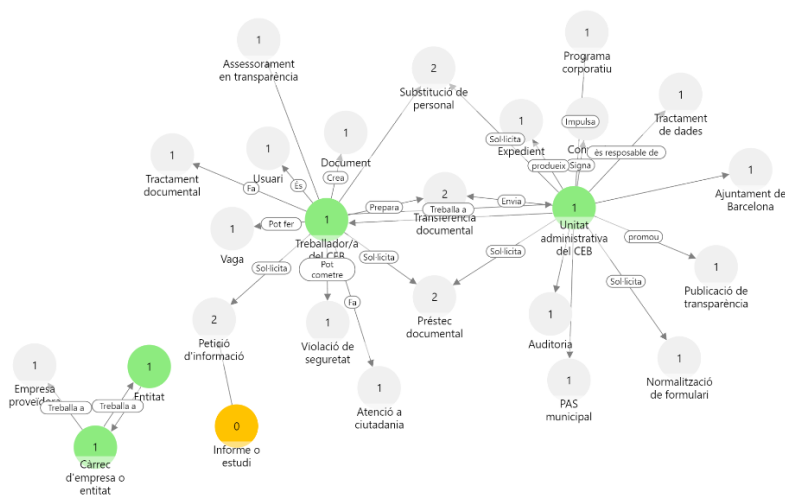
La mitjana per al grup focal dels objectes acció és de 1,69 per sèrie, mentre que per als objectes cosa és de 3,94, amb una mitjana total de 5,63 objectes descrits per sèrie. Per tant, la proporció per a cada sèrie és de 2,33 objectes cosa per cada objecte acció, representant les accions un 30% dels objectes de la sèrie. La variabilitat però és important, com podem observar comparant dues sèries del grup focal, com per exemple la sèrie ORG002-001 Gestió de programes en l'àmbit de processos i sistemes:

Figura 6-1: Mapa d'objectes de la sèrie documental ORG002-001



Amb la sèrie documental de INF001-003 Elaboració d'estudis amb col·laboracions externes:

Figura 6-2: Mapa d'objectes de la sèrie documental INF001-003



SER-09 Nombre de superobjectes

No s'ha aplicat aquest indicador en el cas del Consorci, ja que la poca diferència entre el número d'objectes i superobjectes, i el volum prou abastable d'objectes, ha recomanat basar l'anàlisi en l'indicador anterior.

6.4.4 Indicadors d'anàlisi de les sèries documentals

6.4.4.1 Índex d'influència de la sèrie documental

Atès que en el cas del Consorci d'Educació de Barcelona el nombre de relacions per procés és relativament baix, i es poden establir poques relacions entre sèries (la majoria de sèries només tenen relació amb una altra sèrie), considerem poc significatiu calcular els indicadors SER-10 (índex d'influència simple) i SER-11 (índex d'influència ponderat), atès que ambdós només tenen en compte les relacions entre sèries. Per aquest motiu calculem directament l'indicador SER-12 (índex d'influència net), que té en compte també les relacions entre processos.

SER-12 Índex d'influència net

D'acord amb aquest indicador, les 10 sèries més influents en el conjunt del Consorci són les següents:

Taula 6-18: Sèries documentals amb un major índex d'influència

SÈRIE DOCUMENTAL	ÍNDEX D'INFLUÈNCIA NET
GRH016-002 Jubilacions	28
GRH016-004 Expedients contradictoris	28
GRH016-001 Relació contractual i ocupació de llocs	26
ORG003-001 Anàlisi i millora de processos	26
GRH016-003 Avaluació de personal	24
ECO001-001 Gestió de la despesa	20
INF001-002 Manteniment del banc de dades del CEB	20
NEE008-004 Protocol per avaluar les necessitats de suport [...]	19
ESC005-002 Preinscripció, matrícula i nova admissió	18
GPC005-002 Convocatòria unificada de programes	17

Observem, doncs, que es tracta de sèries clau per al Consorci, que corresponen a tràmits d'inici que donaran origen després a moltes més actuacions, com l'ocupació de llocs de treball (GRH016-001), l'anàlisi i millora de processos (ORG003-001), la preinscripció (ESC005-002) o la convocatòria de programes d'innovació (GPC005-002). Les jubilacions (GRH016-002) i els expedients contradictoris (GRH016-004) es consideren *outliers* a analitzar.

De la mateixa manera, si analitzem les cinc sèries amb un índex més baix:

Taula 6-19: Sèries documentals amb un menor índex d'influència

SÈRIE DOCUMENTAL	ÍNDEX D'INFLUÈNCIA NET
LOP005-001 Auditories APDCAT i controls de seguretat sobre protecció de dades de caràcter personal	-4
NEE010-001 Sol·licituds de preinscripció, reclamacions i nova admissió en centres d'educació especial (CEE)	-3
ORG004-002 Normalització de formularis i models de document	-3
ATU004-001 Llistats de distribució del registre d'entrada	-2
ATU004-002 Albarans d'entrega de correspondència a l'empresa distribuïdora del correu	-2

Observem que es tracta de sèries amb menys rellevància o criticitat sobre l'activitat del Consorci, o bé sèries finalistes que requereixen nombroses actuacions anteriors, com les reclamacions d'escolarització en CEE (NEE010-001) o les auditories de protecció de dades (LOP005-001).

6.4.4.2 Índex de dadificació de la sèrie documental

SER-13 Índex de dadificació simple a nivell d'evidència

De les 189 evidències contingudes en les 32 sèries del grup focal, 86 són interfícies d'interacció amb sistemes d'informació que generen dades estructurades. L'índex mitjà de dadificació, si agrupem les sèries per famílies de processos, és tanmateix de 0,54, cosa la qual indica la presència dins aquestes famílies d'un nombre considerable de sèries altament dadificades. Les mètriques són les següents:

Taula 6-20: Índex de dadificació simple a nivell d'evidència per família de processos

FAMÍLIA	SÈRIES	EVIDÈNCIES DADES	TOTAL EVIDÈNCIES	ÍNDEX DE DADIFICACIÓ ²⁵⁷
Anàlisi i producció d'informació	7	14	24	0,70
Assegurament compliment RGPD	6	12	31	0,39
Atenció a les necessitats educatives especials	3	10	15	0,78
Atenció a les obligacions de transparència	4	12	30	0,59
Escolarització	1	2	10	0,20
Gestió dels recursos humans	2	8	9	0,83
Gestió documental i arxiu	5	18	44	0,43
Racionalització i millora organitzativa	4	10	26	0,38

²⁵⁷ El promig per família no computa les sèries que tenen valor 0 en aquest indicador.

SER-14 Índex de dadificació simple a nivell de dada elemental

De les 596 dades elementals que apareixen en les evidències de les 32 sèries documentals del grup focal, 452 apareixen en evidències de tipus dada. D'acord amb aquest segon indicador, l'índex mitjà de dadificació de les sèries del grup focal creix fins un 0,71. Les mètriques agregades a nivell de família de processos són les següents:

Taula 6-21: Índex de dadificació simple a nivell de dada elemental per família de processos

FAMÍLIA	SÈRIES	DADES ESTRUCTURADES	TOTAL DADES	ÍNDEX DE DADIFICACIÓ ²⁵⁸
Anàlisi i producció d'informació	7	239	268	0,88
Assegurament compliment RGPD	6	50	139	0,35
Atenció a les necessitats educatives especials	3	89	95	0,95
Atenció a les obligacions de transparència	4	41	66	0,79
Escolarització	1	22	36	0,61
Gestió dels recursos humans	2	33	42	0,74
Gestió documental i arxiu	5	72	91	0,79
Racionalització i millora organitzativa	4	35	65	0,56

Observem que en sis de les vuit famílies de processos l'índex de dadificació augmenta amb la segona fórmula de càlcul.²⁵⁹

SER-15 Índex de dadificació ponderat

Els resultats del recàlcul de l'índex de dadificació de la sèrie documental d'acord amb aquesta tercera formulació són els que s'exposen a continuació.

Taula 6-22: Índex de dadificació ponderat per família de processos

FAMÍLIA	SÈRIES	AUT. EN DAD.	AUT. EN DOC.	IMP. EN DAD.	IMP. EN DOC.	ÍNDEX ²⁶⁰
Anàlisi i producció d'informació	7	263	46	47	31	0,50
Assegurament compliment RGPD	6	45	141	25	20	0,27
Atenció a les necessitats educatives especials	3	164	34	18	21	0,48
Atenció a les obligacions de transparència	4	22	43	28	100	0,44
Escolarització	1	22	76	14	0	0,22
Gestió dels recursos humans	2	36	9	33	0	0,54
Gestió documental i arxiu	5	43	43	103	76	0,60
Racionalització i millora organitzativa	4	49	66	62	11	0,32

²⁵⁸ La mitjana per família no computa les sèries que tenen valor 0 en aquest indicador.

²⁵⁹ Una opció, que no hem contemplat en aquest model, és evolucionar aquest indicador fent constar les dades que **només** apareixen a evidències tipus dada, i no a evidències tipus document.

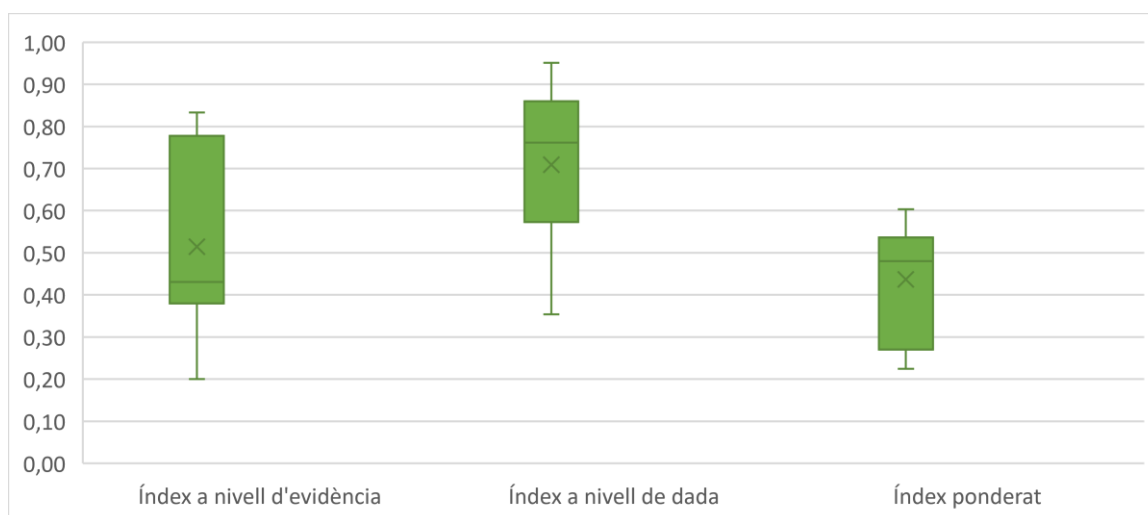
²⁶⁰ La mitjana per família no computa les sèries que tenen valor 0 en aquest indicador.

Si comparem les mesures obtingudes a través de les tres modalitats de càlcul de l'índex de dadificació de la sèrie documental, i ho desagreguem a nivell de sèrie documental del grup focal, obtenim la següent comparativa:

Taula 6-23: Comparativa dels valors dels índex de dadificació per família de processos

FAMÍLIA	ÍNDEX A NIVELL D'EVIDÈNCIA	ÍNDEX A NIVELL DE DADA	ÍNDEX PONDERAT
Anàlisi i producció d'informació	0,70	0,88	0,50
Assegurament compliment RGPD	0,39	0,35	0,27
Atenció a les necessitats educatives especials	0,78	0,95	0,48
Atenció a les obligacions de transparència	0,59	0,79	0,44
Escolarització	0,20	0,61	0,22
Gestió dels recursos humans	0,83	0,74	0,54
Gestió documental i arxiu	0,43	0,79	0,60
Racionalització i millora organitzativa	0,38	0,56	0,32

Gràfic 6-6: Rangos de valors dels tres índex de dadificació



S'observa doncs que l'índex simple a nivell de dada elemental mostra majoritàriament valors superior a l'índex simple a nivell d'evidència, atès que els sistemes d'informació (i les evidències tipus "dades" que se'n deriven) habitualment concentren un major nombre de dades que les evidències-objecte. Però si introduïm mitjançant ponderació la variable del nivell d'integració entre les evidències pel que fa a la compartició de les dades, que és al capdavant l'objectiu principal de la dadificació, observem com el nivell de dadificació real dels processos es redueix notablement, i es fa més homogeni entre tots els processos. Considerem doncs aquesta tercera modalitat (SER-15) com la més representativa del nivell de dadificació òptim d'una sèrie documental, atès que combina la quantitat d'informació estructurada amb el nivell d'integració o de dada única.

6.4.4.3 Índex de ròssec

SER-16 Índex de ròssec intrasèrie

L'índex de ròssec dins les sèries del grup focal és el següent:

Taula 6-24: Índex de ròssec intrasèrie per família de processos

FAMÍLIA	ÍNDEX DE RÒSSEC INTRASÈRIE ²⁶¹
Anàlisi i producció d'informació	0,41
Assegurament compliment RGPD	0,23
Atenció a les necessitats educatives especials	0,19
Atenció a les obligacions de transparència	0,63
Escolarització	0,13
Gestió dels recursos humans	0,57
Gestió documental i arxiu	0,67
Racionalització i millora organitzativa	0,36

L'índex de ròssec de la sèrie documental reflecteix el nivell de reutilització de les dades entre les evidències de la sèrie, cosa la qual és possible només quan el procés està automatitzat i basat en aplicacions integrades entre elles. Es converteix doncs en un indicador d'integració de la informació, i també del nivell d'automatització dels processos.

Amb les dades analitzades observem un nivell baix de reutilització, i per tant d'automatització, en les famílies de processos clau (Escolarització o Atenció a les necessitats educatives especials), que depenen fortament de sistemes externs, mentre aquelles famílies de processos que s'han pogut automatitzar de forma interna, sense dependència de sistemes externs, han pogut assolir un nivell d'integració més alt (Anàlisi i producció d'informació, Atenció a les obligacions de transparència o Gestió documental i arxiu).

SER-17 Índex de ròssec intersèrie

Hem posat a prova l'índex de ròssec entre sèries documentals per avaluar si pot ser un predictor de relacions entre sèries. Analitzant els cinc índex de ròssec intersèrie més alts, observem que es tracta d'un indicador qüestionable.

²⁶¹ El promig per família no computa les sèries que tenen valor 0 en aquest indicador.

Taula 6-25: Relacions entre sèries documentals amb un índex superior de ròssec intersèrie

SÈRIE 1	ÍNDEX	SÈRIE 2
DOC001-002 Assessorament als centres i serveis educatius en matèria de gestió documental	0,75	INF009-002 Manteniment de les dades de la Base de dades de Centres (BDC)
DOC002-001 Tractament documental	0,73	DOC004-001 Préstecs i consultes de documentació
DOC003-001 Transferències a l'Arxiu Central	0,73	DOC004-001 Préstecs i consultes de documentació
DOC002-001 Tractament documental	0,67	DOC003-001 Transferències a l'Arxiu Central
GRH023-001 Gestió de borses de personal de suport als centres educatius (COVID-19)	0,60	INF002-001 Eleccions al Consell Escolar de Centre

Si ens fixem en les relacions estrictament funcionals entre sèries documentals, aquelles detallades en l'apartat 3.6.2.4 de la **GMAP** (capítol 3), no observem coherència en aquest indicador. En alguns casos la coincidència en dades reflecteix relacions reals entre sèries, o entre els processos que generen les sèries, com en totes les sèries de gestió documental (processos amb prefix DOC). Es tractaria en aquest cas de relacions de continuïtat o dependència. Però en d'altres la coincidència en dades no reflecteix cap relació funcional, com en el cas de la Gestió de borses de personal (GRH023-001) amb les Eleccions al Consell Escolar (INF002-001).

Però això no obsta a plantejar la possible utilitat d'aquest indicador per a detectar altres tipus de relacions entre sèries, no contemplades en el nostre model metodològic. Es tractaria de les relacions estructurals, i especialment de les relacions de contingut, que s'expliquen en l'apartat 2.3.1.4.2 del capítol 2. El factor que determina aquesta relació és el fet que les sèries comparteixen descripció sobre els mateixos objectes, amb la qual cosa podem considerar aquest indicador com un element a considerar en el desenvolupament de la conclusió de l'objectiu 7 (apartat 8.1.7 i 8.3.7 del capítol 8).

6.4.4.4 Índex de transaccionalitat

SER-18 Índex de transaccionalitat a partir de les dades elementals

L'índex de transaccionalitat per famílies de les sèries del grup focal, atenent a les dades que descriuen objectes de tipus "Acció", es distribueix de la forma següent:

Taula 6-26: Índex de transaccionalitat a partir de les dades elementals per família de processos

FAMÍLIA	DADES ACCIÓ	TOTAL DADES	ÍNDEX
Anàlisi i producció d'informació	79	268	0,61
Assegurament compliment RGPD	90	139	0,67
Atenció a les necessitats educatives especials	35	95	0,40
Atenció a les obligacions de transparència	36	66	0,43
Escolarització	11	36	0,31
Gestió dels recursos humans	16	42	0,40
Gestió documental i arxiu	46	91	0,53
Racionalització i millora organitzativa	43	65	0,66

SER-19 Índex de transaccionalitat a partir del tipus de dades

Si tenim en compte només el tipus de dada, observem variacions en l'índex de transaccionalitat de les sèries del grup focal:

Taula 6-27: Índex de transaccionalitat a partir del tipus de dades per família de processos

FAMÍLIA	DADES TRANSACCIONALS	TOTAL DADES	ÍNDEX
Anàlisi i producció d'informació	35	268	0,16
Assegurament compliment RGPD	33	139	0,25
Atenció a les necessitats educatives especials	17	95	0,22
Atenció a les obligacions de transparència	15	66	0,21
Escolarització	8	36	0,22
Gestió dels recursos humans	7	42	0,16
Gestió documental i arxiu	25	91	0,32
Racionalització i millora organitzativa	21	65	0,31

La descripció d'una acció pot requerir dades de diferents tipus (geogràfiques, descriptives, etc.) a més de dades transaccionals. Per aquest motiu, l'índex de transaccionalitat basat en el tipus de dades és força més baix que el que té en compte el tipus d'objecte descrit.

SER-20 Índex de transaccionalitat a partir dels objectes

El volum de dades que descriuen cada objecte pot distorsionar el valor de l'indicador SER-18. Per aquest motiu, analitzem directament el nombre d'objectes descrits en el grup focal per donar més precisió a la resta d'indicadors:

Taula 6-28: Índex de transaccionalitat a partir dels objectes per família de processos

FAMILIA	OBJECTES ACCIÓ	TOTAL OBJECTES	ÍNDEX
Anàlisi i producció d'informació	8	40	0,29
Assegurament compliment RGPD	10	28	0,38
Atenció a les necessitats educatives especials	6	20	0,30
Atenció a les obligacions de transparència	6	19	0,34
Escolarització	3	10	0,27
Gestió dels recursos humans	4	11	0,34
Gestió documental i arxiu	12	36	0,32
Racionalització i millora organitzativa	5	15	0,33

Comparem a continuació els tres indicadors per determinar el rang en què es disposen els valors, i observar els valors més freqüents en les sèries del Consorci.

Gràfic 6-7: Rang de valors dels tres índex de transaccionalitat



6.5 Indicadors sobre els processos

6.5.1 Mètriques descriptives dels processos

PRO-01 Nombre de processos

El Consorci d'Educació de Barcelona té actualment actius un total de **217** processos. Aquests processos es distribueixen, segons els tipus clàssics de procés, de la següent forma:

Taula 6-29: Nombre de processos per tipus de procés

TIPUS	PROCESSOS
Estratègic	18
Clau	85
De suport	114

Aquests processos s'han agrupat en 25 famílies de processos, amb la següent quantitat de processos per família:

Taula 6-30: Nombre de processos per família de processos

FAMILIA	PROCESSOS
Anàlisi i producció d'informació	9
Assegurament compliment RGPD	6
Assessorament als centres educatius	10
Atenció a les necessitats educatives especials	19
Atenció a les obligacions de transparència	4
Atenció a l'usuari	5
Comunicació externa	8
Comunicació interna	1
Contractació Administrativa	4
Coordinació de Serveis Educatius	4
Escolarització	7
Gestió de beques i ajuts	7
Gestió de centres	20
Gestió de centres concertats	4
Gestió de les TIC	12
Gestió de programes als centres	11
Gestió de qualificacions	4
Gestió dels recursos humans	24
Gestió documental i arxiu	7
Gestió econòmica	8
Gestió protocol·lària	4

FAMILIA	PROCESSOS
Manteniment i adequació d'equipaments	11
Planificació estratègica	11
Procediments administratius i judicials	10
Racionalització i millora organitzativa	7

D'aquestes, les famílies que tenen un percentatge de processos clau superior al 50%, i per tant les que concentren l'activitat de negoci del Consorci, són les següents:

Taula 6-31: Famílies amb un percentatge més alt de processos clau

FAMILIA	% DE PROCESSOS CLAU
Gestió de beques i ajuts	100%
Atenció a l'usuari	100%
Gestió de centres concertats	100%
Gestió de qualificacions	100%
Atenció a les necessitats educatives especials	90%
Escolarització	86%
Gestió de centres	78%
Coordinació de Serveis Educatius	75%
Gestió de programes als centres	64%

Observem que el major percentatge es concentra en aquelles famílies que tenen com a destinatari a la ciutadania (alumnat i famílies), i el percentatge baixa progressivament en aquelles famílies amb processos que tenen com a destinataris als centres educatius.

6.5.2 Mètriques derivades de les relacions del procés amb altres entitats

PRO-02 Nombre de tasques

El buidatge de les tasques dels processos no ha pogut ser exhaustiu. Per aquest motiu, les magnituds que s'exposen en aquest apartat són parcials, i no reflecteixen el volum real de tasques per procés.

Els casos on el buidatge de tasques ha pogut ser més exhaustiu (els quals s'han assenyalat en negreta en la taula següent) permeten una primera lectura aproximada de la complexitat estimada dels processos a partir del nombre de tasques, que és molt més elevat en processos crítics com els d'Escolarització i d'Atenció a les necessitats educatives especials que no pas en processos interns com l'anàlisi i producció d'informació o la gestió documental i l'arxiu.

Taula 6-32: Percentatge de tasques automatitzades sobre el total de tasques per família de processos

FAMÍLIA DE PROCESSOS	AUTOM.	MANUALS	TOTAL	% AUTOM.
Anàlisi i producció d'informació	6	17	23	26%
Atenció a les necessitats educatives especials	119	149	268	44%
Atenció a les obligacions de transparència	35	5	40	88%
Atenció a l'usuari	7	24	31	23%
Escolarització	48	38	86	56%
Gestió de beques i ajuts	2	35	37	5%
Gestió de centres	0	8	8	0%
Gestió de les TIC	36	18	54	67%
Gestió dels recursos humans	8	16	24	33%
Gestió documental i arxiu	16	17	33	48%
Planificació estratègica	12	2	14	86%
Racionalització i millora organitzativa	7	31	38	18%

Si diferenciem entre tasca, esdeveniment i bifurcació (*gateways*) dels processos, obtenim una segona mesura de la complexitat dels processos del Consorci. Tal com s'ha explicat en el capítol 5, mesurem la complexitat a partir del nombre de bifurcacions o *gateways*, i la dependència de factors externs a través del nombre d'esdeveniments. D'acord amb això, observem que, en el cas del Consorci, la complexitat i la dependència externa dels processos no responen a la importància o criticitat dels processos per a l'organització, sinó a especificitats de cada procés. Per exemple, els processos d'escolarització són processos reglats, on les decisions venen pautades per la normativa, i per tant són menys complexes que els processos de transparència, on cal interpretar la normativa i on les decisions venen condicionades per molts factors que cal valorar de forma independent.

Taula 6-33: Percentatges de complexitat i dependència per família de processos

FAMÍLIA	TASQUES	ESDEVEN.	GATEWAYS	TOTAL	COMPLEXITAT	DEPENDÈNCIA
Anàlisi i producció d'informació	13	6	4	23	17%	26%
Atenció a les necessitats educatives especials	199	42	27	268	10%	16%

FAMÍLIA	TASQUES	ESDEVEN.	GATEWAYS	TOTAL	COMPLEXITAT	DEPENDÈNCIA
Atenció a les obligacions de transparència	24	5	11	40	28%	13%
Atenció a l'usuari	20	4	7	31	23%	13%
Escolarització	66	9	11	86	13%	10%
Gestió de beques i ajuts	29	8	0	37	0%	22%
Gestió de centres	4	4	0	8	0%	50%
Gestió de les TIC	33	15	6	54	11%	28%
Gestió dels recursos humans	14	10	0	24	0%	42%
Gestió documental i arxiu	26	4	3	33	9%	12%
Planificació estratègica	9	3	2	14	14%	21%
Racionalització i millora organitzativa	26	8	4	38	11%	21%

PRO-03 Nombre d'evidències

Les evidències no ha pogut ser informades de forma homogènia en tots els processos. Per aquest motiu, prenem com a mostra les evidències informades en el grup focal, per observar que la mitjana d'evidències per procés té una alta variabilitat en funció del tipus de procés (estratègic, clau i de suport).

Taula 6-34: Nombre i mitjana d'evidències per família de processos

FAMÍLIA DE PROCESSOS	EVIDÈNCIES	MITJANA
Estratègic	63	7,88
Clau	541	15,91
De Suport	247	8,23

PRO-04 Nombre de sèries documentals

El nombre de sèries documentals per procés al Consorci d'Educació de Barcelona són les següents:

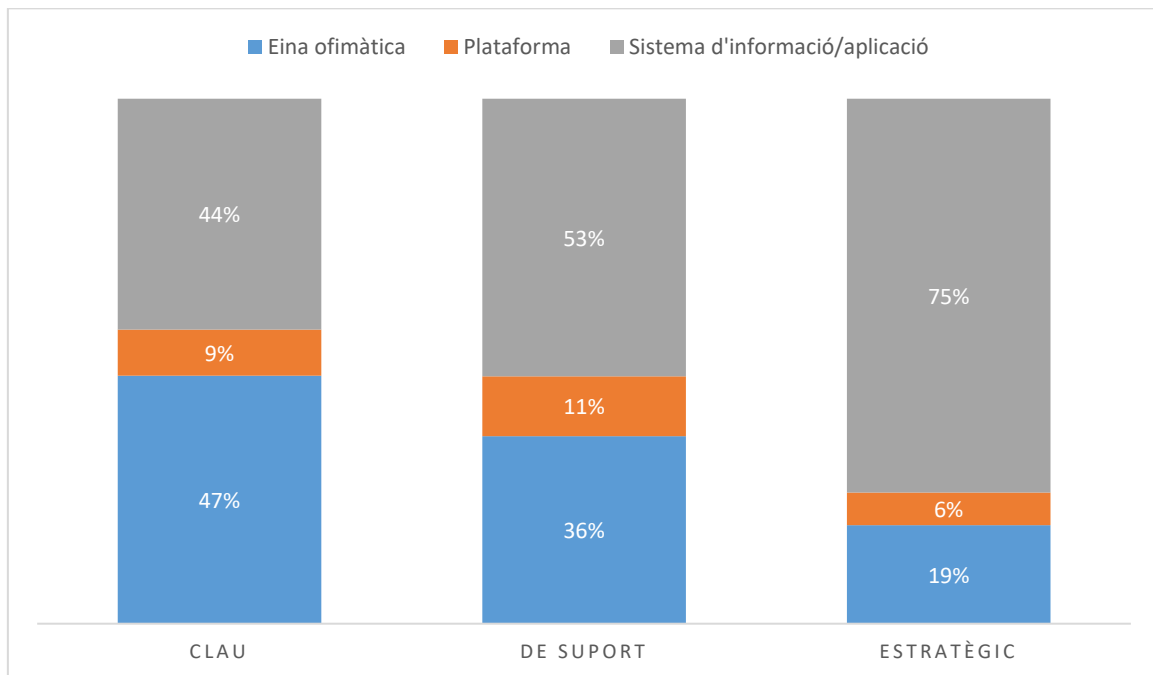
Taula 6-35: Nombre i mitjana de sèries documentals per procés i família de processos

FAMÍLIA DE PROCESSOS	SÈRIES PRIMÀRIES	SÈRIES SECUNDÀRIES	TOTAL SÈRIES	MITJANA
Estratègic	21	4	25	1,39
Clau	127	10	137	1,61
De Suport	190	27	217	1,90

PRO-05 Nombre de sistemes d'informació

Observem els tipus de tecnologia més utilitzats en funció del tipus de procés:

Gràfic 6-8: Entorns tecnològics per tipus de procés

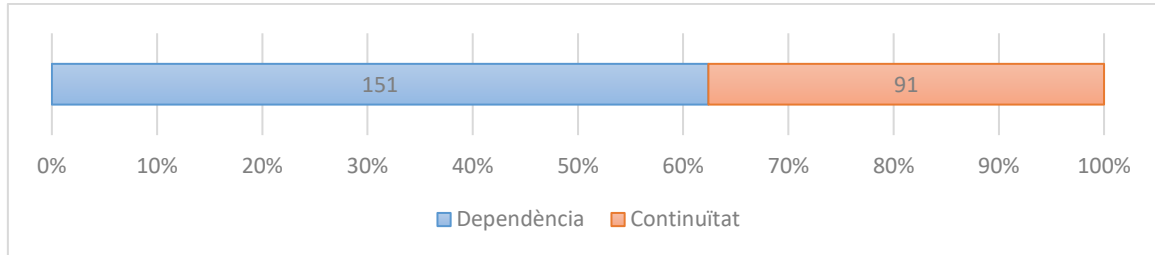


Observem que els processos clau, habitualment processos reglats, utilitzen solucions ofimàtiques en un nombre més elevat que els processos estratègics, que són processos de tipus tècnic. Aquest fenomen, contradictori amb la naturalesa dels processos, s'explica per un motiu d'oportunitat: la transformació digital ha pogut avançar més ràpidament en els processos estratègics i de suport que en els processos clau. Els processos clau són habitualment processos massius, complexes i amb impacte extern, motiu pel qual la substitució d'una eina ofimàtica per una aplicació requereix un desenvolupament important i un sistema d'autenticació per a l'accés a l'aplicació des de fora de l'organització. En canvi, els processos estratègics i de suport tenen requeriments menys exigents, tant a nivell d'autenticació com de procediment, i un impacte extern molt menor, per la qual cosa són candidats a transformacions basades en solucions *low code*, les quals, com veurem en els indicadors sobre sistemes d'informació, han estat la solució predominant en el projecte.

PRO-06 Nombre de processos relacionats

Les mètriques dels tipus de relació entre processos són les següents:

Gràfic 6-9: Tipus de relació entre processos



Observem doncs que les relacions majoritàries són les de dependència, el que implica execucions en paral·lel dels processos, aquelles que determinen menys l'execució del procés.

6.5.3 Mètriques sobre les dades del procés

PRO-07 Nombre de dades elementals

PRO-08 Nombre d'ocurrències de dades

PRO-09 Nombre d'objectes

PRO-10 Nombre de superobjectes

Representem de forma conjunta les mètriques relacionades amb les dades dels 27 processos del grup focal:

Taula 6-36: Nombre d'entitats de dades per procés del grup focal

PROCÉS	DADES	OCURRÈNCIES	OBJECTES	SUPER OBJECTES
TRA004 Atenció a les peticions d'accés a documents i dades del CEB	28	87	6	5
TRA003 Assessorament en matèria de transparència als centres educatius	8	16	5	4
TRA002 Atenció a les peticions d'accés a documents i dades del CEB de la part interessada	26	85	6	5
TRA001 Publicació d'informació a l'espai de transparència	4	5	2	2
ORG004 Normalització de formularis i models de documents	30	115	4	4
ORG003 Anàlisi i millora de processos	16	38	4	4
ORG002 Gestió de programes en l'àmbit de processos i sistemes	19	32	7	3
NEE021 Reconeixement d'alumnat amb NEE-B	45	85	7	7

PROCÉS	DADES	OCURRÈNCIES	OBJECTES	SUPER OBJECTES
NEE020 Intervencions en matèria de seguretat ciutadana i prevenció de radicalismes violents	18	27	5	3
NEE002 Actuacions en casos de sospita de maltractament i/o assetjament en un centre [...]	32	134	8	7
LOP007 Assessorament en matèria de protecció de dades personals	20	25	7	6
LOP006 Manteniment i publicació del registre d'activitats de tractament	10	19	3	3
LOP005 Avaluació i control de la protecció de dades (auditories)	27	67	4	3
LOP003 Gestió de les incidències de seguretat en protecció de dades	40	64	6	3
LOP002 Atenció al drets d'autodeterminació informativa reconeguts per la normativa [...]	20	30	4	4
LOP001 Desenvolupament del sistema de protecció de dades de caràcter personal	22	26	4	4
INF009 Manteniment de les dades de la Base de Dades de Centres (BDC)	101	134	7	4
INF007 Coordinació de l'ús social de centres	28	34	9	5
INF002 Comunicació de les eleccions dels Consells Escolars	39	48	7	3
INF001 Gestió, manteniment i ús del banc de dades del CEB	100	191	17	11
GRH023 Gestió de borses de personal de suport als centres educatius (COVID-19)	25	58	4	3
GRH012 Gestió de les vagues (relacions laborals)	17	20	7	7
ESC001 Seguiment de falsos empadronaments	36	111	10	8
DOC004 Atenció a préstecs i consultes de documentació	33	83	9	9
DOC003 Transferències a l'Arxiu Central	22	74	12	10
DOC002 Tractament documental	15	46	10	10
DOC001 Desenvolupament del sistema de gestió documental	21	62	5	5

6.5.4 Indicadors d'anàlisi dels processos

6.5.4.1 Índex d'influència del procés

PRO-11 Índex d'influència simple
PRO-12 Índex d'influència ponderat
PRO-13 Índex d'influència net

Els índex d'influència dels processos del grup focal que tenen identificades relacions amb altres processos són els següents:

Taula 6-37: Índex d'influència per processos del grup focal

PROCÉS	ÍNDEX SIMPLE	ÍNDEX PONDERAT	ÍNDEX NET
TRA002 Atenció a les peticions d'accés a documents i dades del CEB de la part interessada	1	1	2
TRA001 Publicació d'informació a l'espai de transparència	2	-1	1
ORG004 Normalització de formularis i models de documents	2	-4	-2
ORG003 Anàlisi i millora de processos	10	16	26
ORG002 Gestió de programes en l'àmbit de processos i sistemes	4	7	11
NEE021 Reconeixement d'alumnat amb NEE-B	3	4	7
LOP005 Avaluació i control de la protecció de dades (auditories)	4	-8	-4
LOP003 Gestió de les incidències de seguretat en protecció de dades	1	-2	-1
LOP002 Atenció al drets d'autodeterminació informativa reconeguts per la normativa de protecció de dades personals	1	-2	-1
LOP001 Desenvolupament del sistema de protecció de dades de caràcter personal	4	4	8
INF009 Manteniment de les dades de la Base de Dades de Centres (BDC)	6	-8	-2
INF001 Gestió, manteniment i ús del banc de dades del CEB	1	2	3
ESC001 Seguiment de falsos empadronaments	1	-2	-1
DOC004 Atenció a préstecs i consultes de documentació	1	-2	-1
DOC003 Transferències a l'Arxiu Central	3	2	5
DOC001 Desenvolupament del sistema de gestió documental	2	3	5

6.5.4.2 Índex de dadificació del procés

PRO-14 Índex de dadificació simple a nivell d'evidència
PRO-15 Índex de dadificació simple a nivell de dada elemental
PRO-16 Índex de dadificació ponderat

Per a l'avaluació dels índex de dadificació dels processos del grup focal remetem a les mètriques de dadificació a nivell de sèrie documental, tal com es descriuen a l'apartat 6.4.4.2 d'aquest capítol.

6.5.4.3 Índex de ròssec

PRO-17 Índex de ròssec intraprocés
PRO-18 Índex de ròssec interprocés

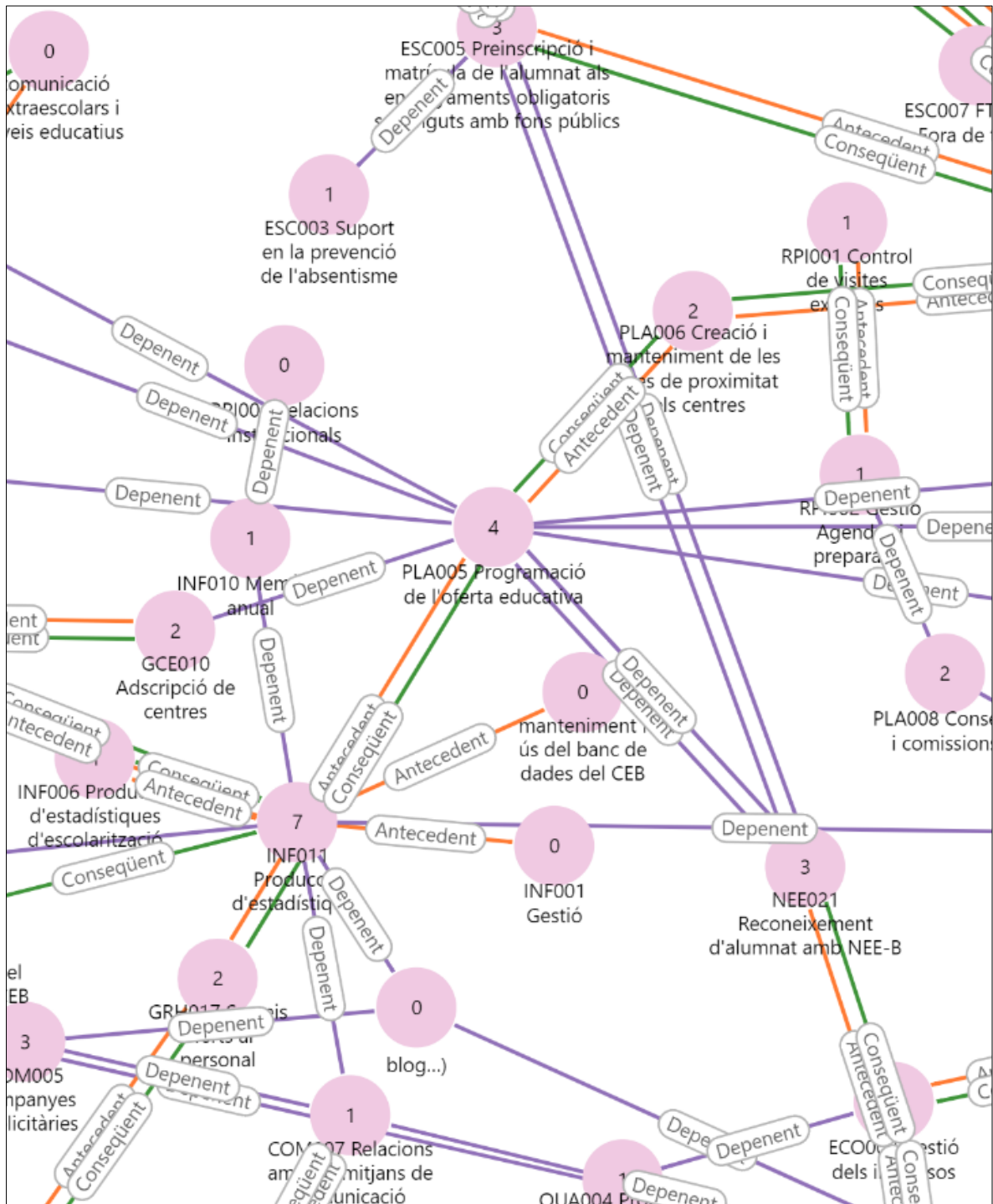
Per a l'avaluació dels índex de ròssec dels processos del grup focal remetem també a les mètriques a nivell de sèrie documental, tal com es descriuen a l'apartat 6.4.4.3 d'aquest capítol.

Extrapolem també les consideracions sobre el possible caràcter predictor de relacions de l'índex de ròssec interprocés, ateses les similituds entre els tipus de relació funcional que poden establir els processos amb els processos (apartat 3.2.4.4 del capítol 3) i les sèries amb les sèries (apartat 3.2.6.4 del capítol 3).

6.5.5 Mapa de les relacions entre processos

El mapa de processos permet representar les relacions entre els processos i el tipus de relació, facilitant la identificació visual dels processos més influents sobre el conjunt del mapa.

Figura 6-4: Fragment del mapa de processos del Consorci d'Educació de Barcelona



6.6 Indicadors sobre els sistemes d'informació

6.6.1 Mètriques descriptives dels sistemes d'informació

SIS-01 Nombre de sistemes d'informació

Actualment el Consorci d'Educació de Barcelona disposa de 451 sistemes d'informació o entorns de treball operatius, que donen cobertura a les necessitat d'informació del Consorci. Aquests sistemes es poden agrupar bàsicament en tres entorns:

Taula 6-38: Nombre de sistemes d'informació per entorn

ENTORN	SISTEMES
Eina ofimàtica	184
Plataforma	37
Sistema d'informació/aplicació	230

Un 41% de tots els sistemes d'informació del Consorci estan basats en eines d'oficina (ofimàtica), un percentatge inferior al que existia durant l'inici del projecte. Si no tenim en compte les plataformes, que es consignen com a entorns de treball però no com a magatzems de dades, constatem que un 44% de la informació estructurada del Consorci s'emmagatzema en solucions ofimàtiques.

Taula 6-39: Nombre de sistemes d'informació per abast

ABAST	SISTEMES
Sistemes gestionats pel Consorci	346
Desenvolupament propi (CEB)	158
Desenvolupament empresa externa (CEB)	14
Desenvolupament usuari final (CEB) ²⁶²	174
Sistemes gestionats per altres ens	105
Departament d'Educació	60
Generalitat de Catalunya	30
Ajuntament de Barcelona	12
Administració General de l'Estat	3

²⁶² En en marc d'aquesta tesi entenem per desenvolupament d'usuari final aquell que té per objectiu satisfer les necessitats funcionals d'un sol usuari o d'un grup reduït d'usuaris, i que és mantingut pels propis usuaris. En el cas del Consorci, s'han desenvolupat algunes solucions utilitzant eines ofimàtiques (10 solucions), però que tenen un abast funcional molt més ampli i un manteniment per part de l'àmbit de tecnologies. En aquest cas s'han computat com a desenvolupaments propis del CEB.

Observem que el Consorci d'Educació de Barcelona només té control directe sobre un 77% dels sistemes d'informació que utilitza. El 23% restant és gestionat per altres administracions. Si obviem la consideració de la ofimàtica com a solució corporativa, el control del Consorci sobre els seus sistemes d'informació es redueix al 62%.

El Consorci, tot i tenir una doble titularitat (Generalitat de Catalunya i Ajuntament de Barcelona), està integrat, als efectes de gestió i contractació, en el model de gestió de les Tecnologies de la Informació de la Generalitat de Catalunya (CTTI). Per aquest motiu observem que el principal proveïdor de sistemes d'informació i aplicacions fora del Consorci és el Departament d'Educació de la Generalitat de Catalunya (52 aplicacions), mentre que el principal proveïdor de plataformes de treball (bàsicament solucions transversals d'Administració electrònica) és la Generalitat de Catalunya (12 plataformes), a través del CTTI. En total, un 86% dels sistemes d'informació que utilitza el Consorci però no gestiona directament pertanyen a la Generalitat de Catalunya.

El model de servei de les solucions es distribueix de la forma següent:

Taula 6-40: Nombre de sistemes d'informació per model de servei

MODEL DE SERVEI	SISTEMES
Adapt. sol. mercat en servidor CTTI	1
Adapt. sol. mercat en servidor extern	1
Desenv. a mida en servidor altra Adm.	6
Desenv. a mida en servidor CTTI	76
Desenv. a mida en servidor extern	4
Desenv. a mida en servidor propi CEB	22
Solucions en xarxa local	183
<i>Software as a Service</i> (corporatiu)	120
<i>Software as a Service</i> (privat)	5
<i>Software as a Service</i> (propi CEB)	24
Solució de mercat en servidor altra Adm.	2
Solució de mercat en servidor CTTI	1
Solució de mercat en servidor extern	3
Solució de mercat en servidor propi CEB	3

Una lectura agregada de les diferents modalitats de servei ens posa de relleu algunes tendències rellevants en l'evolució actual de l'arquitectura de sistemes d'una organització:

Taula 6-41: Lectura agregada dels models de servei

MODEL DE SERVEI AGREGAT	SISTEMES	% TOTAL
Desenvolupaments a mida	108	24%
Solucions en xarxa local	183	41%
Solucions <i>SaaS</i> (<i>cloud</i>)	149	33%
Solucions de mercat	11	2%

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

Observem com, en relació amb la mesura anterior, on un 45% dels sistemes no implicaven desenvolupament, aquest percentatge s'incrementa fins un 57% amb els sistemes que no inclouen cap repositori o base de dades específica per a les dades, i que generen conjunts de dades simples (*dataset*). Aquesta diferència l'explica la utilització intensiva de formularis, que tot i contenir regles de negoci i elements de programació (*low code*) que poden arribar a ser notablement sofisticats, no generen models de dades prou complexes com per requerir una estructura de base de dades. La capacitat d'explotació d'aquests conjunts de dades en un *data lake*, sense la necessitat d'una formalització prèvia a l'explotació (model ELT), permet treballar amb models de dades plans i més simples, i centrar l'esforç de desenvolupament en la interacció amb l'usuari.

Aquest és un canvi de paradigma en el desenvolupament de l'arquitectura de sistemes d'una organització que té importants implicacions en la correspondència entre els conjunts de dades i les agrupacions documentals arxivístiques. Aquest plantejament es desenvolupa a la conclusió de l'objectiu 4 (apartat 8.1.4 del capítol 8).

I una segona observació és el fet que gairebé una quarta part dels sistemes que contenen dades estructurades ho fan en bases de dades orientades a objectes, és a dir una tecnologia especialment orientada a la gestió d'objectes digitals, que ha estat la base de la majoria de solucions EDMS. I això és possible per l'ús que es fa de la capacitat d'aquesta tecnologia de gestionar metadades i assignar-les a objectes, i gestionar en molts casos objectes buits que estan compostats únicament de metadades. Aquesta evidència obre un camp important de reflexió pel que fa a les possibilitats d'aplicació dels conceptes de la gestió documental a la gestió d'informació estructurada.

Una darrera mètrica a considerar és la que correspon a la capacitat dels sistemes d'informació de garantir el valor evidencial de la informació que contenen. Aquest valor ve determinat tècnicament per dues mesures de seguretat: el control de l'accés al sistema, i de l'ús de la seva informació, per part exclusivament d'usuaris autoritzats, i la traçabilitat per part del sistema dels canvis realitzats sobre la informació. Les mesures globals per als sistemes del Consorci són aquestes:

Taula 6-47: Nombre de sistemes d'informació per nivell evidencial del contingut

	Amb traçabilitat	Sense traçabilitat	Total
Amb control d'accés	197	75	272
Sense control d'accés	3	176	179
Total	200	251	451

Aquesta mesura comptabilitza únicament els sistemes de control d'accés i traçabilitat inherents al propi sistema d'informació. No inclou d'altres sistemes perimetrals, com l'autenticació i auditoria dels accessos a xarxa, o els permisos assignats a nivell de carpeta de xarxa local. S'observa clarament que la major part dels sistemes que inclouen control d'accés també garanteixen algun nivell de traçabilitat, mentre que la ofimàtica en

general, representada per aquests 176 sistemes, delega l'autenticació del contingut als sistemes de seguretat perimetrals.

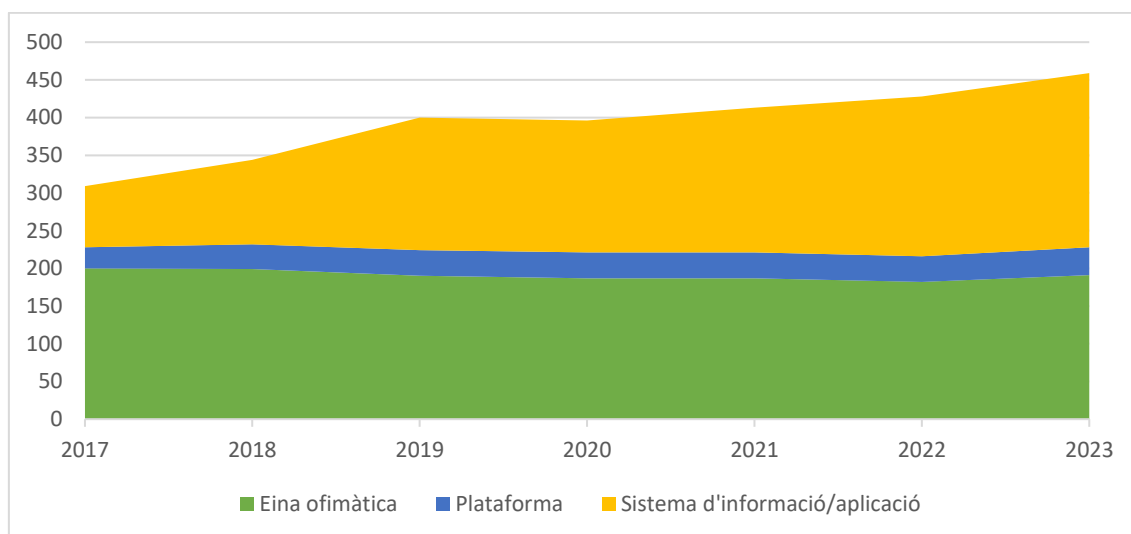
6.6.2 Mètriques de la transformació

Com s'ha comentat a la introducció, el projecte de recollida de dades ha anat acompanyat d'una transformació dels sistemes d'informació del Consorci, adreçada a facilitar la seva dadificació en el grau més alt possible. Per aquest motiu, considerem rellevant descriure amb un conjunt de mètriques com s'ha desenvolupat aquesta transformació fins als 451 sistemes actuals, quin era el punt de partida i com ha afectat aquest procés a les característiques dels sistemes.

SIS-02 Nombre de sistemes d'informació vigents

Es representa a continuació l'evolució en el nombre de sistemes vigents per entorn:

Gràfic 6-14: Evolució del nombre de sistemes d'informació per entorn

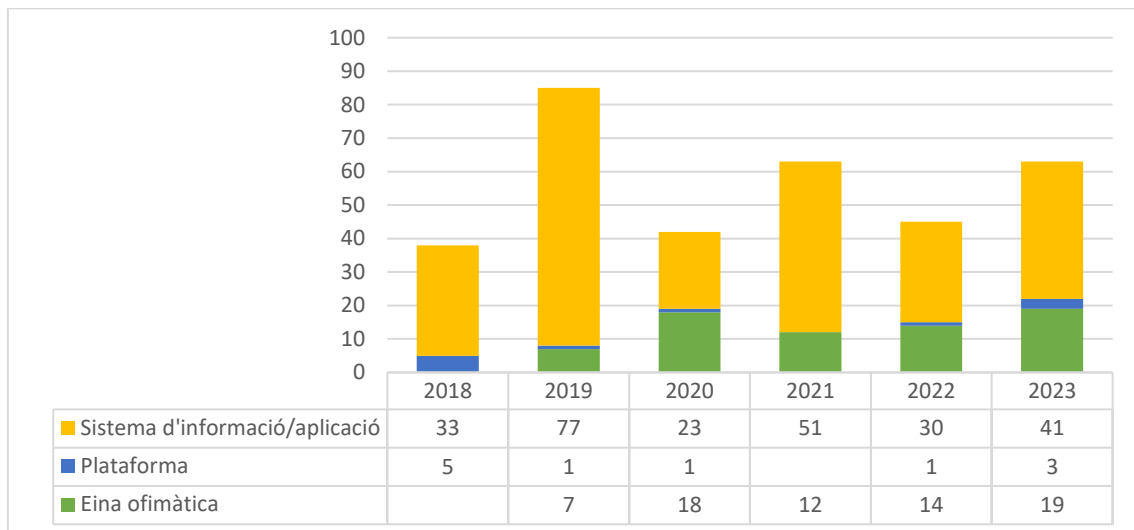


S'observa en el gràfic un creixement sostingut dels sistemes d'informació i aplicacions, propiciat bàsicament per la substitució de solucions ofimàtiques per solucions basades en desenvolupaments *low code*. Això no obstant, no s'aprecia una reducció significativa de les solucions ofimàtiques, tot i la reducció del seu pes sobre el total inicial de sistemes del Consorci. Aquest fenomen es deu a la generació continuada de conjunts de dades simples com a suport a la gestió, producte sovint de formularis electrònics, però també pel seu potencial d'explotació en una arquitectura *data fabric* com la que s'ha anat implantant al Consorci.

SIS-03 Nombre de sistemes d'informació desenvolupats

Els sistemes d'informació posats en producció per any i entorn són els següents:

Gràfic 6-15: Nombre de sistemes d'informació posats en producció per any i entorn



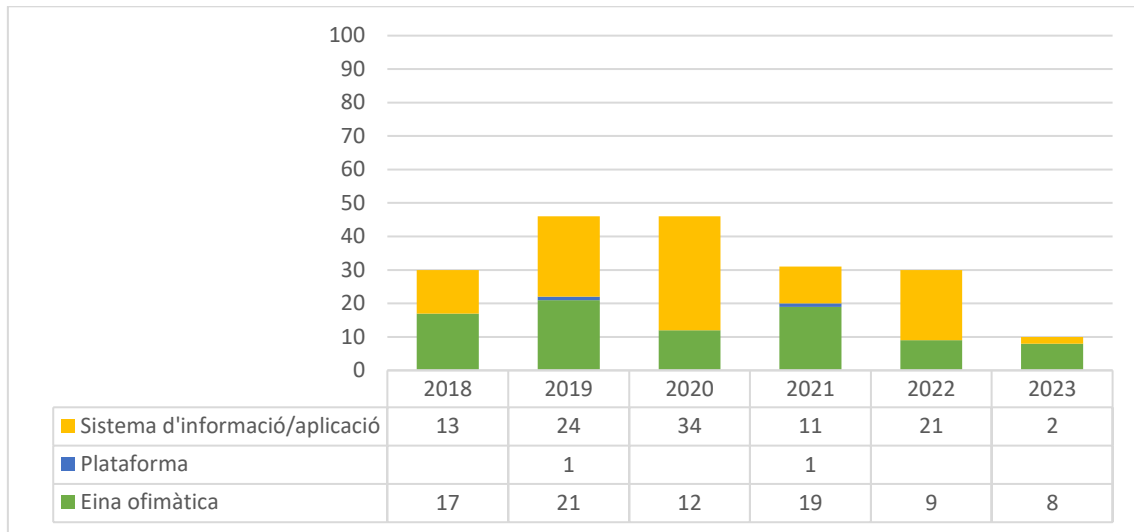
No es comptabilitza l'any 2017 per evitar distorsions, atès que, en desconèixer a l'inici del projecte la data exacta de posada en producció de la major part de sistemes d'informació, tots els sistemes vigents en aquell moment es van atribuir de forma convencional a l'any 2017. Per tant, ubiquem l'inici de la dadificació del Consorci a l'any 2018.

En aquest sentit, s'observa un desenvolupament constant de sistemes d'informació i aplicacions, bàsicament propiciat per l'ús de solucions *low code*.

SIS-04 Nombre de sistemes d'informació desactivats

Pel que fa als sistemes d'informació desactivats per obsolescència funcional, tecnològica, o per substitució, les magnituds per any i entorn són les següents:

Gràfic 6-16: Nombre de sistemes d'informació desactivats per any i entorn



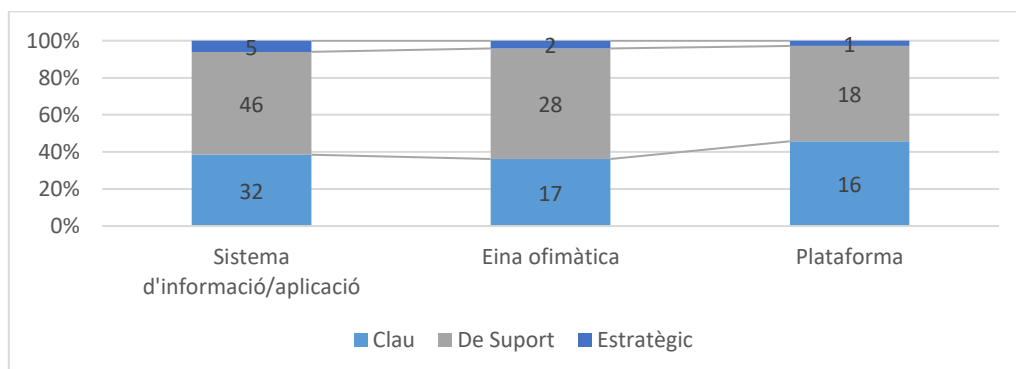
Observem en aquest cas també unes magnituds constants, amb una lleugera reducció de les solucions ofimàtiques però també amb constant caducitat d'aplicacions. En aquest segon cas, aquesta obsolescència obeeix a la vinculació de moltes solucions a convocatòries anuals, especialment en el cas dels formularis electrònics.

6.6.3 Mètriques derivades de la relació del sistema d'informació amb altres entitats

SIS-05 Nombre de processos

Del total de processos del Consorci (217), n'hi ha 118 que, a través del treball de camp, s'han pogut vincular a sistemes d'informació. Cada procés pot utilitzar més d'un sistema, i al mateix temps un sistema pot ser utilitzat per més d'un procés. Per tant, en aquest cas només podem fer una valoració a nivell categòric.

Gràfic 6-17: Nombre de sistemes d'informació per entorn tecnològic i tipus de procés



La proporció en la presència dels diferents tipus de procés en cada entorn tecnològic no varia més enllà de 5 punts percentuals, cosa la qual indica que el tipus de procés no és una variable determinant per a l'ús majoritari d'un determinat entorn tecnològic.

L'anàlisi de les dades desagregades no es pot considerar rellevant pel fet de no haver pogut vincular tots els sistemes a processos existents. Però de les famílies de processos de l'estudi focal, on això ha estat possible, s'ha pogut observar que el nombre de processos als quals dona suport un sistema d'informació pot esdevenir un indicador significatiu de la rellevància d'aquell sistema en el mapa funcional de l'organització. Així doncs, si prenem els tres sistemes que tenen més de quatre processos vinculats:

Taula 6-48: Sistemes d'informació amb més de quatre processos vinculats

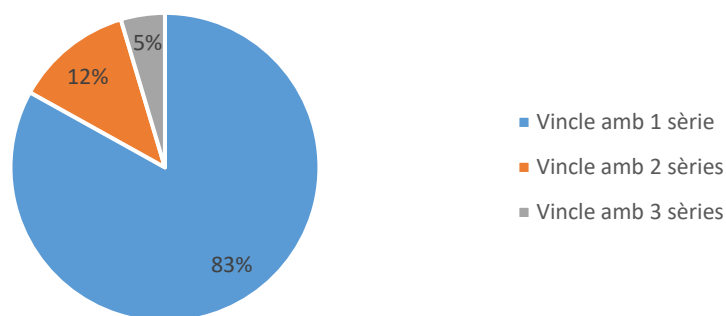
SISTEMA	Processos
SI04.008 GEDAC	9
SI02.012 Catàleg de processos del CEB	5
SI03.002 Web del CEB	5

Observarem que el SI04.008 GEDAC, el sistema que gestiona la preinscripció, i per tant un sistema directament vinculat a la família de processos Escolarització, també està donant suport a processos de cinc famílies més: Atenció a l'usuari, Planificació estratègica, Anàlisi i producció d'informació, Gestió de beques i ajuts, i Gestió de les TIC. Però el segon a la llista, el SI02.012 Catàleg de processos del CEB, concentra en una sola família, Racionalització i millora organitzativa, la major part dels processos vinculats. La rellevància per al conjunt de l'organització del SI04.008 GEDAC serà doncs superior, no només per la magnitud de processos vinculats, sinó per la seva dispersió entre famílies diferents de processos.

SIS-06 Nombre de sèries documentals

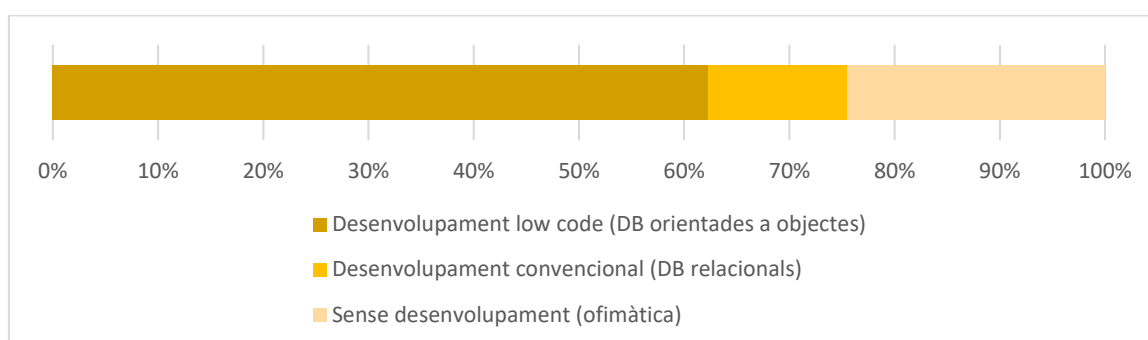
Si analitzem tots els sistemes d'informació que, en el mapa de sistemes, s'han pogut vincular a sèries documentals vigents, observem que, en la majoria dels casos, cada sistema dona suport només a una sèrie documental:

Gràfic 6-18: Percentatge del nombre de vincles de sistemes d'informació amb sèries documentals



El perfil tecnològic dels sistemes d'informació és el següent:

Gràfic 6-19: Perfil tecnològic dels sistemes d'informació vinculats a sèries documentals



Anem a fer una lectura més detallada, però en sentit invers. Si ens centrem en les sèries documentals dels processos del grup focal, podem extreure també dades significatives. De les 32 sèries documentals del grup focal, 26 sèries tenen vincles amb 37 sistemes d'informació, però només en dos casos es tracta d'eines ofimàtiques. La resta són aplicacions

Els entorns tecnològics són els següents:

Taula 6-49: Entorns tecnològics de les sèries documentals del grup focal

ENTORN	MODEL DE DESENVOLUPAMENT	TECNOLOGIA	SISTEMES	SÈRIES
Sistema d'informació/ aplicació	Desenvolupament convencional	[Barra grisa]	1	1
		[Barra grisa]	1	1
		[Barra grisa]	1	3
		[Barra grisa]	2	2

ENTORN	MODEL DE DESENVOLUPAMENT	TECNOLOGIA	SISTEMES	SÈRIES
			7	4
	Desenvolupament low code		18	18
	No identificat		4	4
Eina ofimàtica	Sense desenvolupament		2	2
Plataforma			1	1

Aquest indicador ha estat determinant per a les conclusions dels objectius 3 i 4 (apartats 8.1.3 i 8.1.4 del capítol 8).

SIS-07 Nombre d'evidències

El nombre d'evidències que interaccionen a cada sistema d'informació és un indicador rellevant del nivell d'utilització del sistema. Les dades generals, amb les evidències informades actualment als mapes, i agregades per tecnologia d'emmagatzemament, són les següents:

Taula 6-50: Mitjana d'evidències per entorn tecnològic

ENTORN	MITJANA EVIDÈNCIES PER SISTEMA
Eina ofimàtica	1,13
Plataforma	2,04
Sistema d'informació/aplicació	2,01

Observem que la mitjana d'evidències per sistema és en general baixa: molts sistemes, bàsicament formularis d'entrada de dades o manteniments bàsics d'expedients, només impliquen un o dos accessos durant tot el procés. I això és especialment acusat amb la ofimàtica, on en la majoria dels casos és impossible comptabilitzar en nombre real d'accessos per actualitzar dades, i s'ha optat per comptabilitzar un únic accés tipus per procés o sèrie (equivalent a obertura d'expedient). Els cinc sistemes on s'ha descrit més interaccions d'usuari recollides en forma d'evidències són processos dadificats:

Taula 6-51: Sistemes d'informació amb major nombre d'evidències relacionades

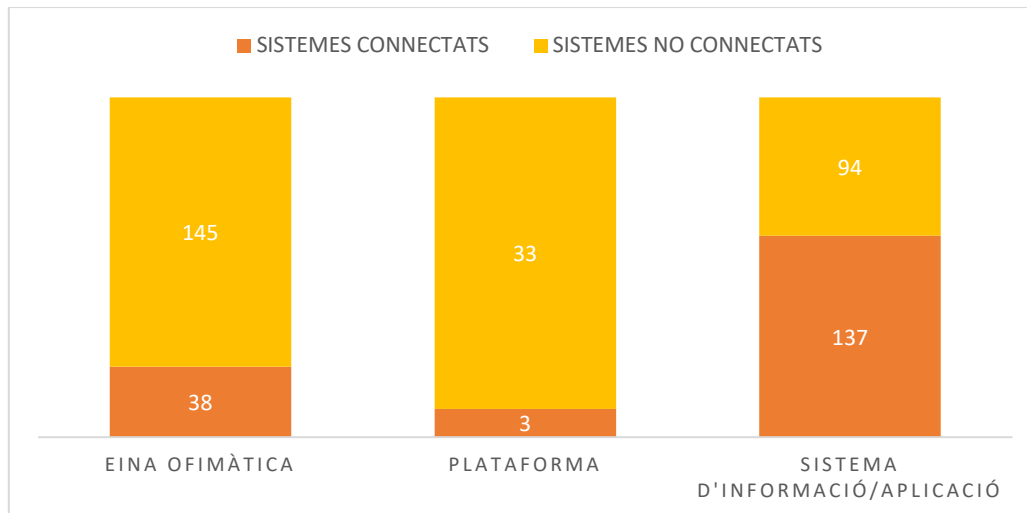
SISTEMA	EVIDÈNCIES
SI05.004 BOGA Gestió de Beques i Ajuts	14
SI02.011 Sistema d'ajuts de menjador	10
SI04.006 GTE Gestió de títols PRE-LOGSE	10
SI08.055 NEE023-001 Alumnes atesos per UAO	9
SI05.034 GEAC (Aplicació de Gestió Expedients de Arxiu Central)	8

SIS-08 Nombre de relacions amb sistemes d'informació

Actualment hi ha actives 202 interfícies de connexió entre sistemes vigents. Aquestes interfícies connecten un total de 178 sistemes d'informació amb 17 sistemes receptors de dades. Aquests 178 sistemes representen un 39% del total de sistemes d'informació vigents del Consorci. Cal tenir present que el 41% dels sistemes d'informació del Consorci són sistemes ofimàtics, és a dir sistemes que en teoria no es connecten a d'altres sistemes. Si no incloem la ofimàtica ni les plataformes, observem que el 77% de totes les aplicacions del Consorci intercanvien dades amb algun altre sistema d'informació.

Però la connectivitat de la ofimàtica s'incrementa de forma notable quan es treballa amb un model *data lake* o *data fabric*, com s'aprecia en el següent gràfic:

Gràfic 6-20: Nombre de sistemes d'informació connectats a altres sistemes per entorn tecnològic



Observem també en l'anàlisi que els 178 sistemes d'informació proveïdor de dades les envien a només 17 sistemes receptors de dades. Aquesta situació s'explica per l'existència al Consorci d'un gran sistema d'informació que fa les funcions de *data lake*, el qual s'anomena SI02.014 Banc de Dades del Consorci d'Educació de Barcelona (BI-CEB). Dels 178 sistemes d'informació, 168 estan connectats al Banc de Dades. Les

interaccions entre sistemes que no passin pel Banc de Dades del CEB afecten només a 19 sistemes d'informació.

La interconnexió entre sistemes d'informació punt a punt, sense una plataforma que faciliti aquesta connexió, és lenta i costosa per a les organitzacions. A les conclusions es desenvolupa que l'ús de plataformes tipus *data lake* o *data fabric* facilita la connectivitat entre sistemes i accelera la possibilitat d'intercanviar i reutilitzar dades dins de la pròpia organització.

El nombre d'interfícies de connexió en funció del tipus d'interfície i procediment són els següents:

Taula 6-52: Nombre d'interfícies en funció del tipus i procediment

TIPUS INTERFÍCIE	PROCEDIMENT INTERCANVI	INTERFÍCIES
Accés a aplicació	Manual directe	17
	Robotitzat	20
Automatitzada	Connexió directa a origen	93
Obtenció manual de fitxer	Manual amb depuració prèvia	33
	Manual directe	39

Els períodes d'actualització de dades en cada interfície són els següents:

Taula 6-53: Nombre d'interfícies en funció del període d'actualització

PERÍODE D'ACTUALITZACIÓ	INTERFÍCIES
A demanda	48
Anual	25
Diària	45
Mensual	3
Per Esdeveniment	75
Puntual Programat	1
Setmanal	5

Finalment, la generació de la versió històrica o d'arxiu de cada intercanvi de dades, es fa amb les freqüències següents:

Taula 6-54: Nombre d'interfícies en funció del termini de generació de la versió d'arxiu

HISTORIFICACIÓ	INTERFÍCIES
Acumulativa	109
Temporal (estadística)	93

És rellevant observar en aquest cas que més de la meitat dels sistemes no apliquen una política de captura per talls temporals (anuals o per curs acadèmic), sinó que acumulen la informació.

6.6.4 Mètriques sobre les dades del sistema d'informació

SIS-09 Nombre de dades elementals

Durant el treball de camp no ha estat possible completar el buidatge de les dades elementals de tots els sistemes d'informació del Consorci, donada la dificultat per accedir a les interfícies o al model de dades d'alguns sistemes. Dels 451 sistemes d'informació vigents del Consorci, s'ha pogut recollir i informar les dades completes de 202 sistemes (un 45% del total). Les mètriques que es faciliten a continuació provenen doncs d'aquests sistemes.²⁶⁵

El buidatge de les dades ha permès detectar en els 202 sistemes un total de 1647 dades elementals, que es fan presents a través de 2779 ocurrències en sistemes. Aquestes dades es distribueixen, en funció de les característiques dels sistemes com segueix:

Taula 6-55: Dades elementals per entorn dels sistemes d'informació

ENTORN	SISTEMES	DADES ELEMENTALS	MITJANA DE DADES PER SISTEMA	OCURRÈNCIES	MITJANA D'OCURRÈNCIES PER SISTEMA
Eina ofimàtica	98	448	5	815	8
Sistema d'informació/aplicació	104	1288	12	1964	19

Taula 6-56: Dades elementals per abast dels sistemes d'informació

ABAST	SISTEMES	DADES ELEMENTALS	MITJANA DE DADES PER SISTEMA	OCURRÈNCIES	MITJANA D'OCURRÈNCIES PER SISTEMA
CEB empresa externa	1	13	13	13	13
CEB propi	94	900	10	1355	14

²⁶⁵ També s'ha recollit durant el projecte 555 dades de sistemes ja obsolets, que no es tenen en compte per aquesta anàlisi.

ABAST	SISTEMES	DADES ELEMENTALS	MITJANA DE DADES PER SISTEMA	OCURRÈNCIES	MITJANA D'OCURRÈNCIES PER SISTEMA
Departament d'Ensenyament	7	442	63	583	83
Generalitat de Catalunya	1	3	3	3	3
Usuari final	99	452	5	825	8

Taula 6-57: Dades elementals per model d'emmagatzemament dels sistemes d'informació

MODEL D'EMMAGATZEMAMENT	SISTEMES	DADES ELEMENTALS	MITJANA DE DADES PER SISTEMA	OCURRÈNCIES	MITJANA D'OCURRÈNCIES PER SISTEMA
Base de dades orientada a objectes (GD)	52	574	11	745	14
Base de dades relacional	25	589	24	827	33
Conjunts de dades simples	122	606	5	1123	9
No identificat	3	66	22	84	28

Taula 6-58: Dades elementals per model de desenvolupament dels sistemes d'informació

MODEL DE DESENVOLUPAMENT	SISTEMES	DADES ELEMENTALS	MITJANA DE DADES PER SISTEMA	OCURRÈNCIES	MITJANA D'OCURRÈNCIES PER SISTEMA
Desenvolupament convencional	25	589	24	827	33
Desenvolupament low code	90	799	9	1243	14
Sense desenvolupament	84	368	4	625	7
No identificat	3	66	22	84	28

El tipus de dades que conté el sistema d'informació pot ser un element rellevant per classificar el sistema d'informació en funció de les categories que s'indiquen a les conclusions.

Taula 6-59: Nombre de dades elementals per tipus de dada i entorn

TIPUS DE DADA	EN EINA OFIMÀTICA	EN SISTEMA D'INFORMACIÓ/APLICACIÓ
Acadèmiques	25	127
Contacte / Comunicació	24	56
Descriptives	44	265
Econòmiques	39	82
Identificatives (persona i registre)	130	285
Laborals o contractuals	71	107
Localització o geogràfiques	21	86
Sanitàries	7	2
Socials	13	5
Tecnològiques	0	4
Transaccionals / de tramitació	44	237
Volumetries / Consums	30	32

SIS-10 Nombre d'ocurrències de dades

Les mètriques de les variables que apareixen en evidències del tipus "Dades" vinculades a sistemes d'informació són les següents:

Taula 6-60: Nombre d'ocurrències per caràcter de l'ocurrència i entorn

ENTORN	AUTÒCTONES	IMPORTADES	NO INFORMADES
Eina ofimàtica	5	32	0
Plataforma	1	0	8
Sistema d'informació/aplicació	352	321	301

SIS-11 Nombre d'objectes

SIS-12 Nombre de superobjectes

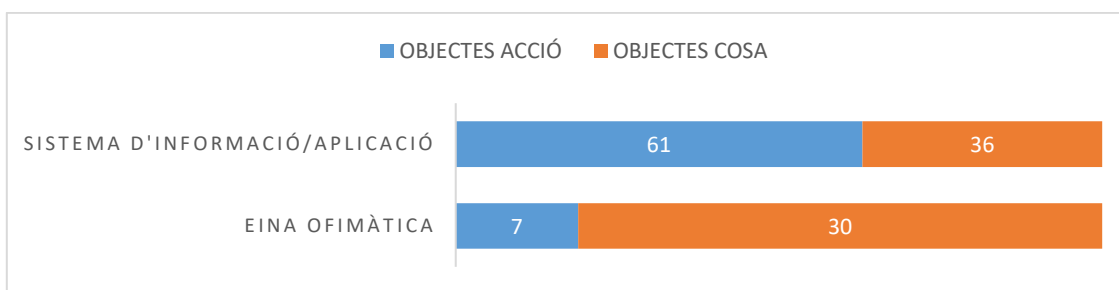
La mitjana d'objectes "acció" per sistema d'informació és de 0,76, mentre que la mitjana d'objectes del tipus "cosa" és de 2,08. La mitjana d'objectes per sistema no arriba a 3 objectes (2,85).

La primera observació és que els sistemes d'informació tenen unes proporcions d'objectes acció i d'objectes cosa molt similars a les sèries documentals. Així doncs, mentre que la proporció cosa/acció en les sèries documentals és de 2,33 "coses" per cada "acció", com ha mostrat l'indicador SER-08 (apartat 6.4.3 d'aquest capítol), en el cas dels sistemes la proporció puja lleugerament, i és de 2,73 "coses" per cada "acció".

En general, els sistemes d'informació tenen una presència major d'objectes tipus "cosa" que les sèries documentals, i en una proporció major. Això es deu a que tota sèrie documental fa referència com a mínim a una acció (amb una mitjana de 1,69 accions per sèrie), mentre que un sistema d'informació pot tenir caràcter registral, ser compartit per diferents sèries documentals, i no fer referència a cap acció (d'aquí l'índex d'accions per sistema inferior a 1).

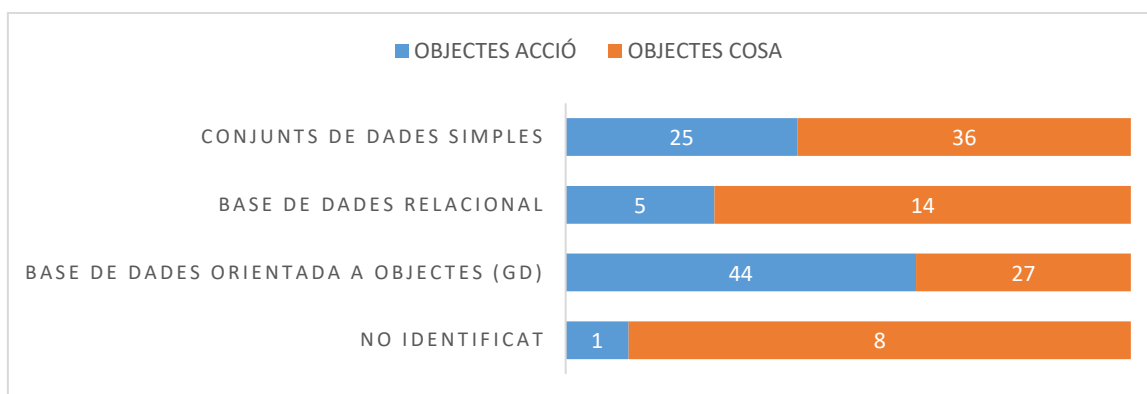
Quan incorporem la variable tecnologia del sistema, observem tanmateix que la proporció canvia. Mentre en les eines ofimàtiques es manté la proporció observada de forma general, i fins i tot s'aguditzza, en les aplicacions, entorn habitual de la tramitació administrativa, la proporció s'inverteix.

Gràfic 6-21: Proporció d'objectes tipus "cosa" i "acció" per entorn tecnològic



Desglossant l'entorn tecnològic pel sistema d'emmagatzemament, observem:

Gràfic 6-22: Proporció d'objectes tipus "cosa" i "acció" per tecnologia d'emmagatzemament



La tecnologia més utilitzada per a gestionar informació d'accions ha estat la que segueix el model d'un gestor documental (base de dades orientada a objectes), és a dir aquella orientada a la gestió d'expedients. La proporció que s'invertia en el gràfic 6-21 s'explica doncs pel fet que, en la transformació digital portada a terme al Consorci, aquesta tecnologia ha estat la base dels desenvolupaments *low code* que han permès substituir massivament la ofimàtica en aquells processos on es requeria un entorn de tramitació o gestió d'expedients. En les conclusions de l'objectiu 4 s'aprofondeix en aquesta qüestió (apartat 8.1.4 del capítol 8).

6.6.5 Indicadors d'anàlisi dels sistemes d'informació

6.6.5.1 Índex d'influència del sistema d'informació

Com s'ha exposat en el capítol 5, aquest indicador es planteja en quatre modalitats.

SIS-13 Índex de criticitat

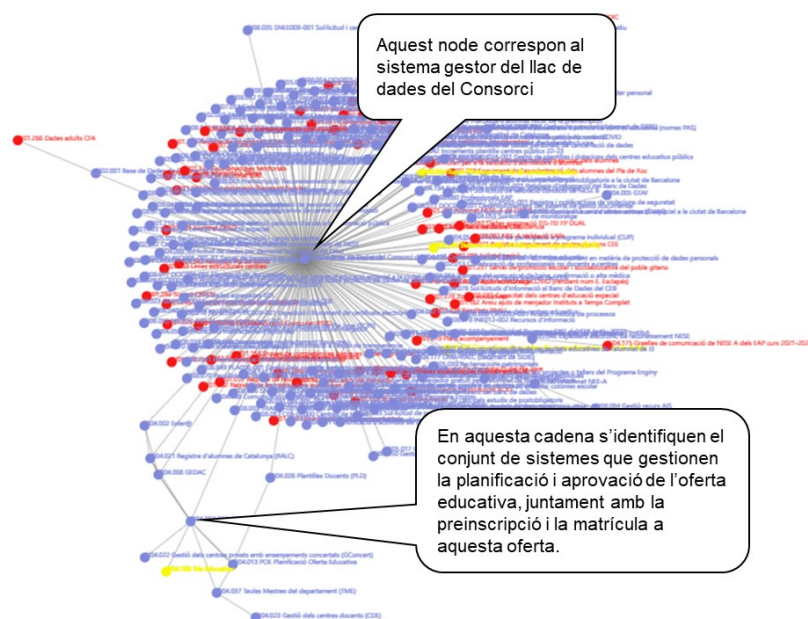
Els sistemes del Consorci que podríem considerar crítics en funció del nombre de relacions que mantenen amb altres sistemes són els següents:

Taula 6-61: Relació de sistemes d'informació amb un índex de criticitat més alt

SISTEMES D'INFORMACIÓ	ÍNDEX DE CRITICITAT
SI02.014 Banc de Dades del Consorci d'Educació de Barcelona (BI-CEB)	173
SI04.012 OFE-EDU	5
SI10.001 Alumnat NEE-A	5
SI04.008 GEDAC	2
SI04.021 Registre d'alumnes de Catalunya (RALC)	2

Aquest indicador també permet una detecció visual a través del mapa de sistemes d'informació, que en el cas del Consorci ens permet detectar els dos grans nuclis de sistemes relacionats:

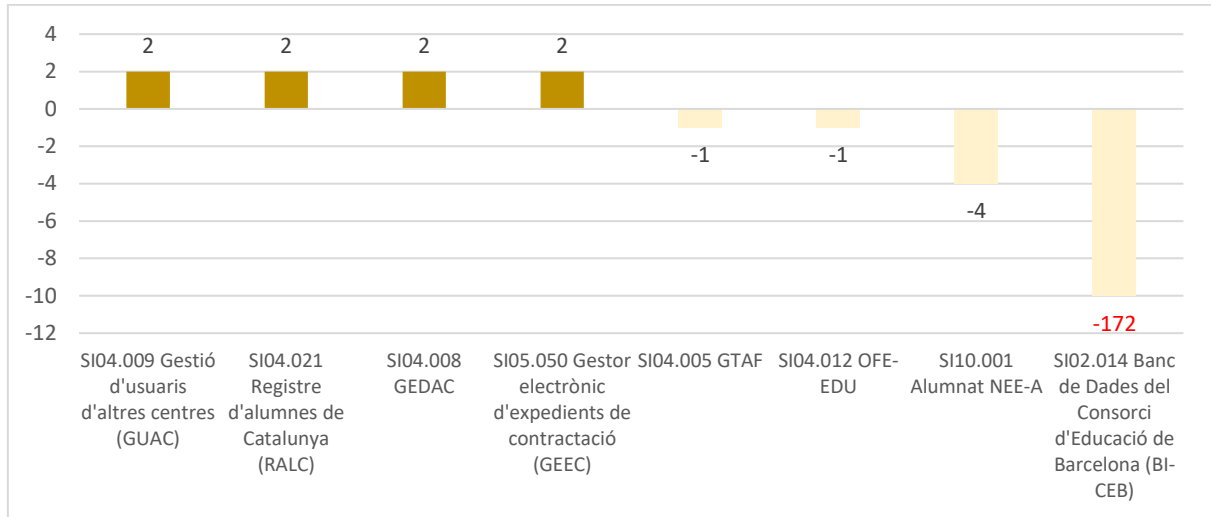
Figura 6-5: Mapa d'interaccions entre sistemes d'informació del Consorci d'Educació de Barcelona



SIS-14 Índex d'influència

Si analitzem quins són els sistemes més influents i més dependents a partir d'aquest indicador, observem el següent:²⁶⁶

Gràfic 6-23: Sistemes d'informació amb major i menor índex d'influència



Analitzant els resultats, observem que la influència d'un sistema no permet cap lectura funcional, atès que la rellevància funcional dels sistemes representats en el diagrama no respon a l'ordre en el qual els disposa aquest indicador. La influència d'un sistema sobre altres sistemes ve determinada, per tant, només per les connexions entre sistemes.

SIS-15 Índex de connectivitat ponderada

Aplicant aquesta ponderació, observem que els 10 sistemes d'informació amb més connectivitat del Consorci són els següents:²⁶⁷

Taula 6-62: Relació dels deu sistemes d'informació amb un índex de connectivitat més alt

SISTEMA D'INFORMACIÓ	TECNOLOGIA	ÍNDEX
SI04.012 OFE-EDU		26
SI04.008 GEDAC		17
SI10.001 Alumnat NEE-A		16
SI04.021 Registre d'alumnes de Catalunya (RALC)		14

²⁶⁶ En la gràfica hem modificat l'aspecte visual del valor del Banc de Dades del CEB per facilitar la llegibilitat de la resta de valors de la gràfica.

²⁶⁷ En aquest cas considerem el Banc de Dades del CEB com un *outlier*, atès que el seu caràcter de *data lake* no el fa comparable a d'altres solucions pel que fa a connectivitat.

SISTEMA D'INFORMACIÓ	TECNOLOGIA	ÍNDEX
SI04.013 POE Planificació Oferta Educativa		12
SI04.037 Taules Mestres del departament (TME)		9
SI04.002 Esfer@		7
SI02.015 Registre d'Alumnat Vulnerable (RAV)		6
SI04.009 Gestió d'usuaris d'altres centres (GUAC)		6
SI08.067 Gestió recurs SIEI		6

Observem doncs que els sistemes d'informació de processos clau basats en tecnologia relacional consolidada són els que apareixen amb un índex de connectivitat més alt.

6.6.5.1.1 Índex de connectivitat potencial

SIS-16 Índex de connectivitat potencial

L'anàlisi de l'índex de connectivitat potencial, és a dir la coincidència entre els diccionaris de dades dels sistemes, no ha donat resultats significatius, probablement pel fet que no s'han pogut buidar les dades de tots els sistemes vigents al Consorci.

6.6.5.2 Índex de transaccionalitat del sistema

Recordem que aquest indicador té dues variants:

SIS-17 Índex de transaccionalitat a partir de les dades elementals

L'índex de transaccionalitat mitjà dels sistemes d'informació vigents del Consorci amb dades elementals informades és d'un 0,34.

El valor d'aquest indicador agregat per les diferents característiques dels sistemes d'informació és el següent:²⁶⁸

²⁶⁸ No es contemplen en aquest cas les plataformes, atès que l'objectiu en aquest cas és comparar tecnologies, i les plataformes no corresponen a tecnologies específiques. De fet, és una categoria que el resultat d'aquesta anàlisi ens porta a replantejar.

Taula 6-63: Índex de transaccionalitat dels sistemes d'informació per entorn

ENTORN	TOTAL DADES	DADES ACCIÓ	ÍNDEX
Eina ofimàtica	448	101	0,23
Sistema d'informació/aplicació	1288	466	0,36

Taula 6-64: Índex de transaccionalitat dels sistemes d'informació per abast

ABAST	TOTAL DADES	DADES ACCIO	INDEX
CEB empresa externa	13	0	0,00
CEB propi	900	399	0,44
Departament d'Ensenyament	442	60	0,14
Generalitat de Catalunya	3	0	0,00
Usuari final	452	105	0,23

Taula 6-65: Índex de transaccionalitat dels sistemes d'informació per model d'emmagatzemament

MODEL D'EMMAGATZEMAMENT	TOTAL DADES	DADES ACCIO	INDEX
Base de dades orientada a objectes (GD)	574	354	0,62
Base de dades relacional	589	72	0,12
Conjunts de dades simples	606	147	0,24
No identificat	66	1	0,02

Taula 6-66: Índex de transaccionalitat dels sistemes d'informació per model de desenvolupament

MODEL DE DESENVOLUPAMENT	TOTAL DADES	DADES ACCIO	INDEX
Desenvolupament convencional	589	72	0,12
Desenvolupament <i>low code</i>	799	418	0,52
Sense desenvolupament	368	81	0,22
No identificat	66	1	0,02

A partir de les quatre agregacions de l'índex de transaccionalitat basat en dades, podem deduir que el perfil tipus d'un sistema d'informació de caràcter transaccional (és a dir, que dona suport a la gestió d'un procés), es caracteritza per ser una aplicació, desenvolupada preferentment per part del Consorci, i majoritàriament basada en sistemes orientats a objectes, tipus gestor documental, parametritzats amb un desenvolupament lleuger (*low code*). Aquest perfil coincideix amb la solució que s'ha utilitzat al Consorci per accelerar la dadificació dels processos, transformant bàsicament solucions preexistents basades en ofimàtica. Si aquest perfil es trasllada a la situació inicial del projecte, es conclou que l'any 2017 el nucli de la informació transaccional del Consorci s'estava gestionant bàsicament amb ofimàtica.

SIS-18 Índex de transaccionalitat a partir del tipus de dades
SIS-19 Índex de transaccionalitat a partir dels objectes

En el cas dels sistemes d'informació s'ha centrat l'anàlisi en l'índex a partir dels objectes. Els valors d'aquest indicador per tipus de tecnologia són els següents:

Taula 6-67: Índex de transaccionalitat basat en objectes per entorn

ENTORN	TOTAL OBJECTES	OBJECTES ACCIÓ	ÍNDEX
Eina ofimàtica	37	7	0,19
Sistema d'informació/aplicació	97	61	0,63

Taula 6-68: Índex de transaccionalitat basat en objectes per model d'emmagatzemament

MODEL D'EMMAGATZEMAMENT	TOTAL OBJECTES	OBJECTES ACCIÓ	ÍNDEX
Base de dades orientada a objectes (GD)	71	44	0,62
Base de dades relacional	19	5	0,26
Conjunts de dades simples	61	25	0,41
No identificat	9	1	0,11

6.6.5.3 Classificació dels sistemes pel nivell de seguretat de les dades

SIS-20 Classificació per nivell de seguretat de les dades

Com s'ha descrit en el capítol 5, en la classificació del nivell de seguretat dels sistemes d'informació s'ha tingut en compte l'entorn tecnològic del sistema, l'abast institucional, i l'existència de controls de la integritat, i de l'accés i l'ús. D'acord amb aquesta classificació obtenim una primera aproximació al nivell de seguretat mitjà per abast, puntuat en una escala de l'1 al 10:

Taula 6-69: Classificació de seguretat dels sistemes d'informació per abast

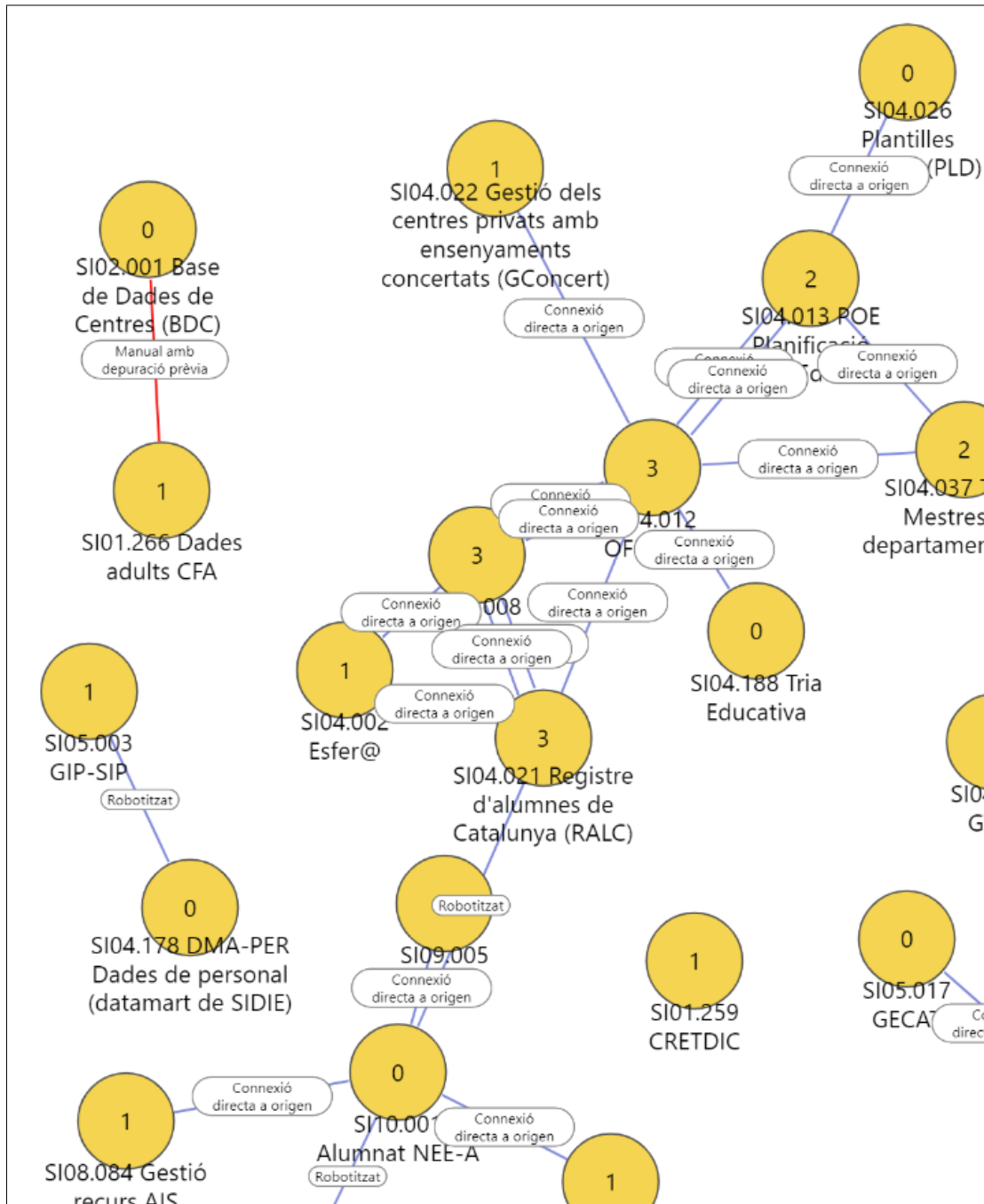
ABAST	SEGURETAT
[Barra horitzontal]	[Barra vertical]
[Barra horitzontal]	[Barra vertical]
[Barra horitzontal]	[Barra vertical]
[Barra horitzontal]	[Barra vertical]
[Barra horitzontal]	[Barra vertical]
[Barra horitzontal]	[Barra vertical]
[Barra horitzontal]	[Barra vertical]

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

6.6.6 Mapa de relacions entre sistemes d'informació

El mapa de relacions entre sistemes visualitza les interaccions entre sistemes, i pot representar les característiques de les interfícies. En l'exemple que es presenta es visualitza un aspecte concret de la interfície, que és el procediment de transferència de les dades entre els sistemes.

Figura 6-6: Fragment del mapa de sistemes d'informació del Consorci d'Educació de Barcelona



6.7 Indicadors sobre les dades

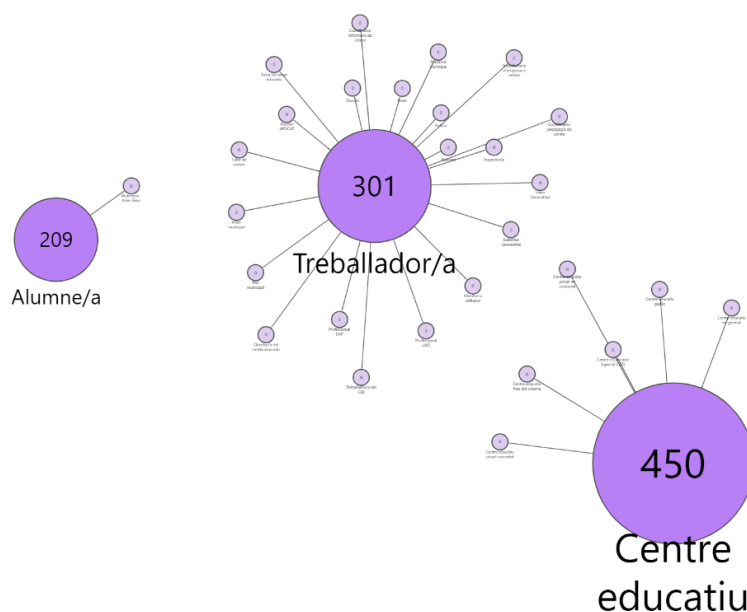
6.7.1 Mètriques descriptives de les dades elementals

DAD-01 Nombre de dades elementals

El mapa de dades del Consorci té catalogades un total de 2291 dades elementals. 977 d'aquestes dades descriuen objectes del tipus “acció”, mentre que 1314 descriuen objectes de tipus “cosa”.

Els superobjectes descrits per un major nombre de dades són, en primer lloc, el “centre educatiu” (450 dades elementals), en segon lloc el “treballador/a”²⁶⁹ (301 dades elementals), i en tercer lloc l’alumne/a (209 dades elementals).

Figura 6-7: Superobjectes descrits per un major nombre de dades elementals

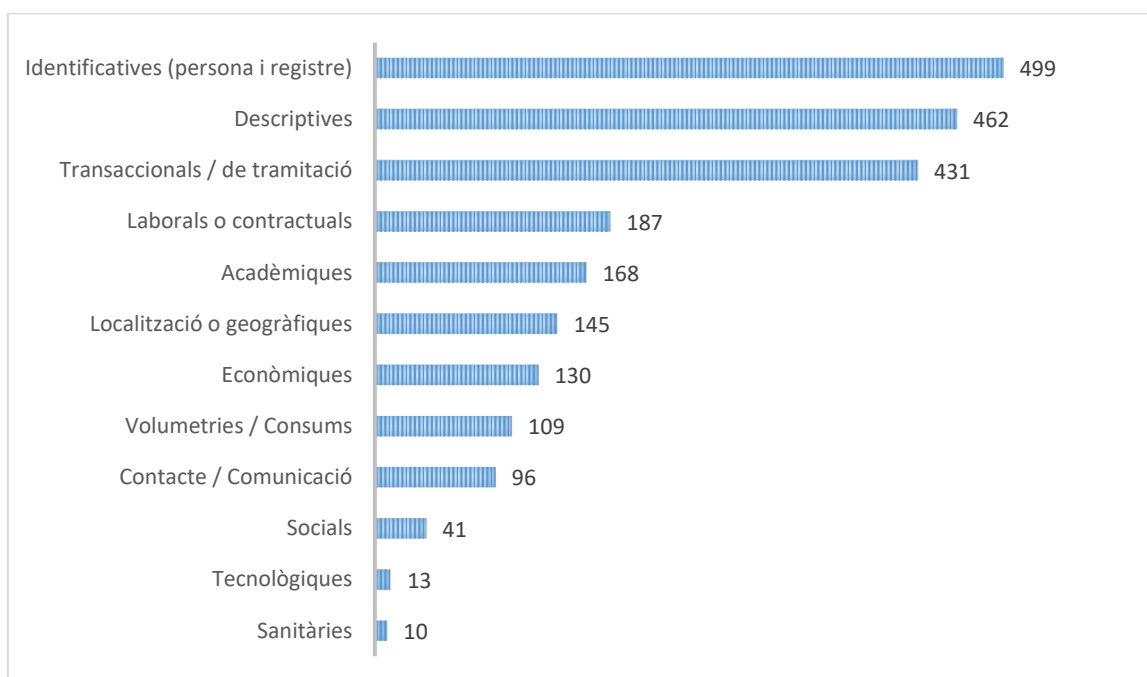


La dada elemental més utilitzada en sistemes d’informació és “Nom de centre” (educatiu), amb 130 ocurrències en sistemes d’informació (102 en sistemes vigents), mentre que la dada amb més ocurrències en evidències és “Nom unitat administrativa”, amb 117 ocurrències.

El nombre de dades elementals per cada tipus de dada es mostra en el següent gràfic:

²⁶⁹ Entès com qualsevol figura professional vinculada a la comunitat educativa.

Taula 6-70: Nombre de dades elementals per tipus de dada



DAD-02 Nombre d'ocurrències de dades

A dia d'avui el Consorci té registrades al mapa global un total de 3334 assignacions de dades a sistemes d'informació, 2779 de les quals són assignacions a sistemes vigents. Pel que fa a les evidències, ha fet 2013 assignacions de dades a evidències, de les quals 2008 són a evidències vigents o vinculades a sistemes vigents.

DAD-03 Nombre d'objectes

Pel que fa als objectes, el mapa de dades del Consorci conté actualment informació sobre un total de 172 objectes, 92 dels quals corresponen a "accions" i 80 a "coses".

DAD-04 Nombre de superobjectes

Els 172 objectes s'agrupen en 67 superobjectes, 42 dels quals són de tipus "acció" i 29 dels quals de tipus "cosa".

Podria sorprendre que el nombre d'objectes i superobjectes de tipus "acció" superi el de tipus "cosa", tenint present que l'indicador SER-08 (apartat 6.4.3 d'aquest capítol) ens aportava la proporció de 2,33 objectes cosa per cada objecte acció recollits en processos i sèries documentals. Però cal tenir present que, mentre que cada procés i

sèrie generen tot sovint un nou objecte de tipus “acció”, en el cas dels objectes “cosa” els processos només recullen informació d’objectes que ja existeixen, i poques vegades en creen de nous. Així doncs, el mapa d’objectes complet d’una organització hauria de tendir a tenir un nombre d’objectes “acció” similar al de processos, mentre que el nombre d’objectes “cosa” serà inferior.

6.7.2 Mètriques derivades de la relació de la dada elemental amb altres entitats

6.7.2.1.1 A nivell de dada elemental

DAD-05 Nombre de sistemes d’informació
DAD-06 Nombre d’evidències
DAD-07 Nombre de processos
DAD-08 Nombre de sèries documentals

Les deu dades elementals amb valors més elevats en els quatre indicadors anteriors són les següents:²⁷⁰

Taula 6-71: Dades elementals amb major presència en processos, sèries documentals, sistemes i evidències

DADA	SISTEMES	DADA	EVIDÈNCIES	PROCESSOS	SÈRIES
Nom de centre educatiu	130	Nom unitat administrativa	117	20	24
Codi de centre educatiu	77	Nom i cognoms treballador CEB	86	21	24
Nom i cognoms de docent	40	Codi de centre educatiu	53	22	10
Nom i cognoms d'alumne/a	38	Nom de centre educatiu	46	21	11
Tipus de centre educatiu	29	Correu electrònic treballador CEB	32	11	14
Nom i cognoms treballador CEB	28	Correu electrònic de centre educatiu	28	9	9
DNI/NIF de Docent	27	IDALU d'alumne/a	26	9	3
IDALU d'alumne/a	27	Nom i cognoms d'alumne/a	22	8	2
Districte de centre educatiu	26	Nom i cognoms de docent	22	7	5
Correu electrònic de centre educatiu	20	Nom i cognoms - Extern ciutadà	21	4	4

²⁷⁰ Cal tenir present que aquests indicadors no compten ocurrences, sinó vincles amb actius d’informació identificats a través de les ocurrences.

6.7.2.1.2 A nivell d'objecte

DAD-09 Nombre de sistemes d'informació
DAD-10 Nombre d'evidències
DAD-11 Nombre de processos
DAD-12 Nombre de sèries documentals

Els deu objectes amb més presència en els quatre indicadors anteriors són els següents:

Taula 6-72: Objectes amb major presència en processos, sèries documentals, sistemes i evidències

OBJECTE	SISTEMES	OBJECTE	EVIDÈNCIES	PROCESSOS	SÈRIES
Centre educatiu en general	152	Unitat administrativa del CEB	118	20	24
Alumne/a	70	Treballador/a del CEB	93	23	25
Docent	44	Centre educatiu en general	65	26	14
Treballador/a del CEB	38	Alumne/a	39	15	4
Servei educatiu en general	32	Docent	24	8	5
Contracte de serveis	24	Ciutadania en general	22	5	4
Família / Tutor/a legal	20	Petició d'accés	21	2	2
Director/a del centre educatiu	19	Família / Tutor/a legal	20	7	3
Unitat administrativa del CEB	19	Atenció a ciutadania	17	3	3
Empresa proveïdora	13	Director/a del centre educatiu	17	9	9

6.7.2.1.3 A nivell de superobjecte

DAD-13 Nombre de sistemes d'informació
DAD-14 Nombre d'evidències
DAD-15 Nombre de processos
DAD-16 Nombre de sèries documentals

Els deu superobjectes amb més presència en els quatre indicadors anteriors són els següents:

Taula 6-73: Superobjectes amb major presència en processos, sèries documentals, sistemes i evidències

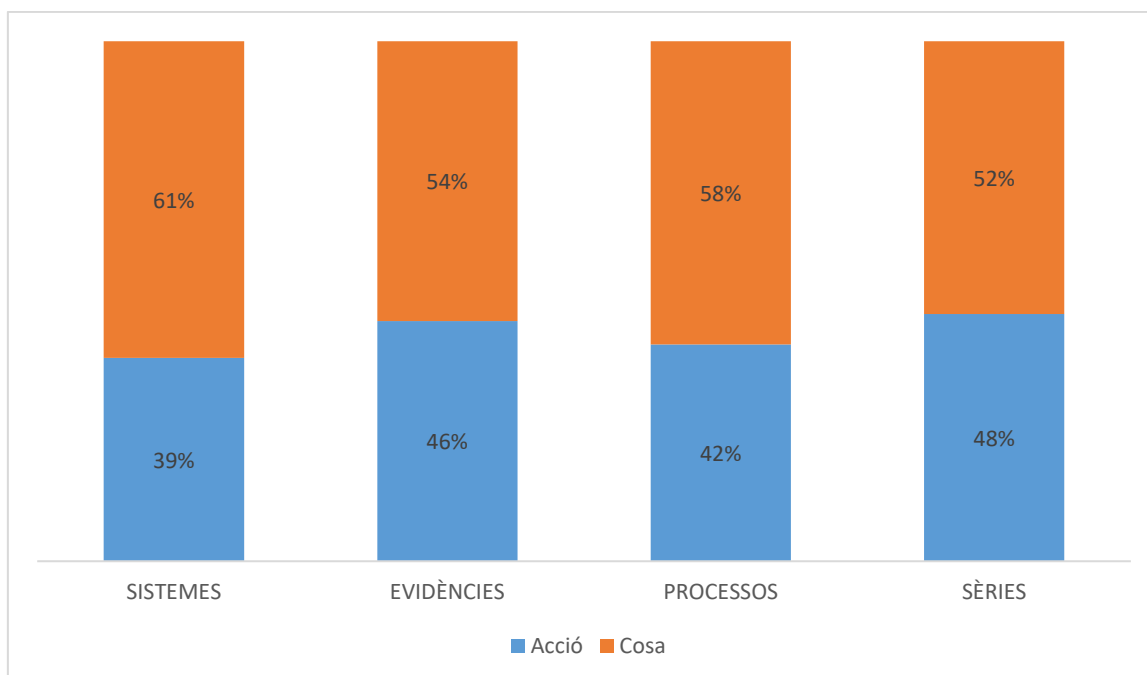
SUPEROBJECTE	SISTEMES	SUPEROBJECTE	EVIDÈNCIES	PROCESSOS	SÈRIES
Centre educatiu	154	Treballador/a	123	28	29
Treballador/a	102	Ens de l'Admin. pública	120	20	25
Alumne/a	73	Centre educatiu	66	27	15
Servei de suport educatiu	43	Alumne/a	39	15	4
Contracte	28	Tractament documental	27	3	3
Ens de l'Admin. pública	24	Ciutadania	22	5	4
Família	20	Atenció i tramitació	20	4	4
Empresa	17	Família	20	7	3
Escolarització	17	Entitat	16	9	6
Entitat	13	Projecte	16	4	3

6.7.2.1.4 A nivell de tipus d'objecte (cosa/acció)

DAD-17 Nombre de sistemes d'informació
DAD-18 Nombre d'evidències
DAD-19 Nombre de processos
DAD-20 Nombre de sèries documentals

El percentatge de presència dels dos tipus d'objecte en els quatre indicadors és la següent:

Gràfic 6-26: Percentatge de dades elementals descriptives de "coses" i "accions" en processos, sèries, sistemes i evidències



Continuant amb el raonament iniciat en el comentari als indicadors DAD-03 i DAD-04, observem que la proporció de dades sobre "accions" en relació amb les dades sobre "coses" és similar a la que mostra l'indicador SER-08: la informació sobre accions que trobem en qualsevol sèrie documental és aproximadament un 40% de tota la informació que conté la sèrie. I la mateixa proporció mitjana observem en aquesta gràfica.

6.7.3 Indicadors d'anàlisi de les dades elementals

6.7.3.1 Índex de rellevància de la dada elemental

DAD-21 Índex de rellevància dins el procés o la sèrie documental
--

DAD-22 Índex de rellevància dins el sistema d'informació
--

En el cas del Consorci només s'ha calculat l'índex de rellevància global.

DAD-23 Índex de rellevància global

D'acord amb aquest índex, les deu dades més rellevants per al conjunt del mapa de dades del Consorci d'Educació de Barcelona són les següents:

Taula 6-74: Índex de rellevància global de les dades elementals del Consorci d'Educació de Barcelona

DADA ELEMENTAL	PROCESSOS	SÈRIES	EVIDÈNCIES	SISTEMES	ÍNDEX
Nom de centre	21	11	46	130	0,18
Nom unitat administrativa	20	24	117	15	0,15
Codi de centre	22	10	53	77	0,14
Nom i cognoms treballador CEB	21	24	86	28	0,13
Nom i cognoms del docent	7	5	22	40	0,06
Nom i cognoms de l'alumne/a	8	2	22	38	0,06
Correu electrònic treballador CEB	11	14	32	11	0,06
Correu electrònic de centre	9	9	28	20	0,06
IDALU d'alumne/a	9	3	26	27	0,05
Nom i cognoms director/a de centre	9	9	17	18	0,04

6.7.3.2 Índex de rellevància de l'objecte i del superobjecte

6.7.3.2.1 Índex de rellevància global

DAD-24 Índex de rellevància global

D'acord amb aquest indicador, els deu objectes més rellevants per al Consorci són els següents:

Taula 6-75: Índex de rellevància global dels objectes del Consorci d'Educació de Barcelona

OBJECTE	ÍNDEX
Centre educatiu en general	0,22
Treballador/a del CEB	0,15
Unitat administrativa del CEB	0,15
Alumne/a	0,11
Docent	0,07
Director/a del centre educatiu	0,05
Família / Tutor/a legal	0,04
Entitat	0,04
Empresa proveïdora	0,03
Ciutadania en general	0,03

DAD-25 Índex de rellevància global

D'acord amb aquest indicador, els cinc superobjectes més rellevants per al Consorci són els següents:

Taula 6-76: Índex de rellevància global dels superobjectes del Consorci d'Educació de Barcelona

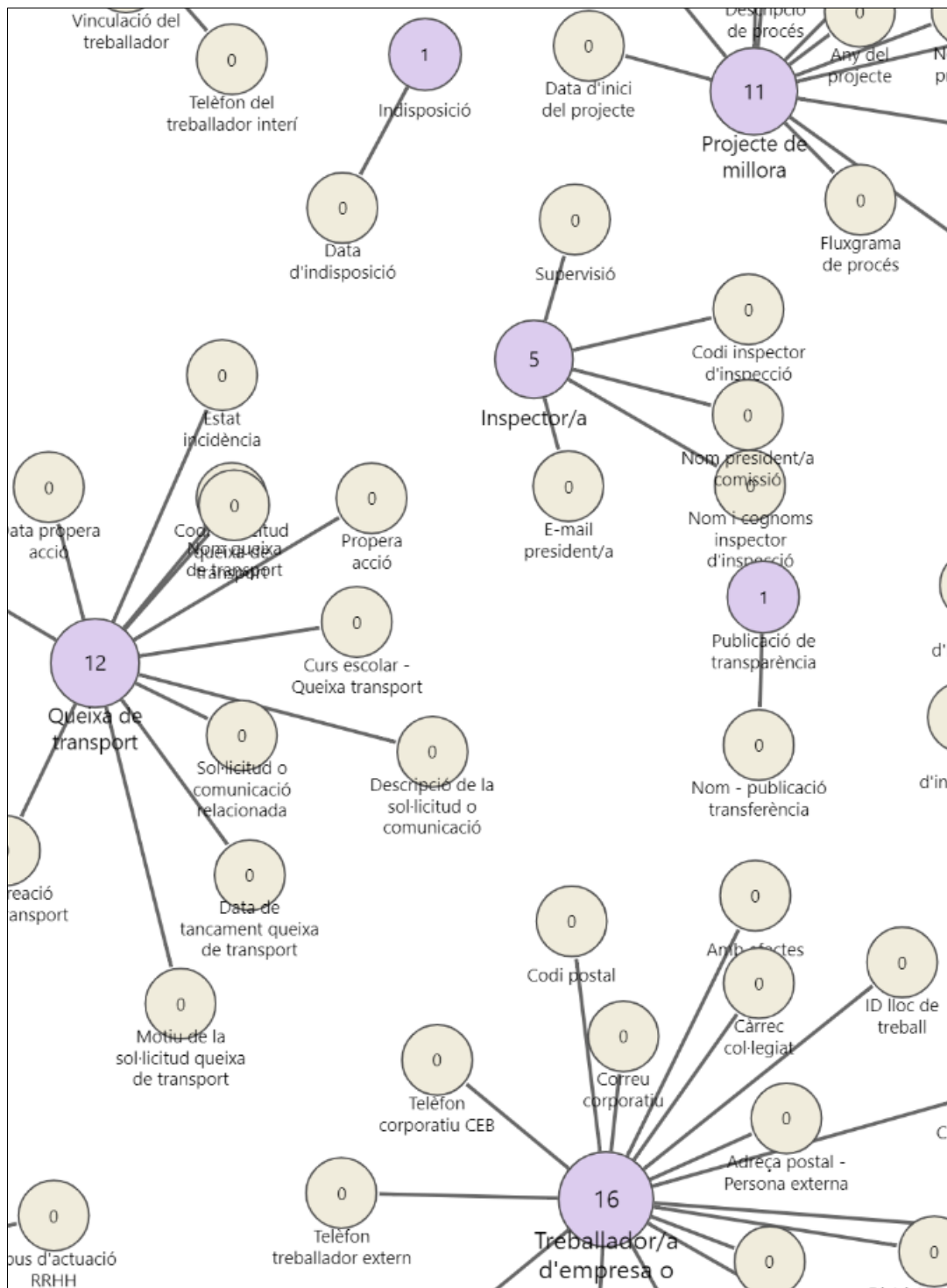
SUPEROBJECTE	ÍNDEX
Treballador/a	0,24
Centre educatiu	0,22
Ens de l'Administració pública	0,16
Alumne/a	0,11
Família	0,04

Aquesta visió és il·lustrativa dels eixos de negoci del Consorci, i de la càrrega de treball administratiu que genera la gestió de cada eix. El dibuix que emergeix d'aquestes dades mostra un escenari en el qual les polítiques del Consorci d'Educació de Barcelona es despleguen dotant de recursos humans els centres educatius per tal de prestar serveis a l'alumnat. Si analitzem el pes en processos i recursos esmerçats, observem que l'estructura orgànica del Consorci (serveis centrals) disposa de tants recursos humans per a la gestió de professionals com per a la gestió directa de centres educatius o serveis a les famílies, cosa la qual coincideix amb l'escenari apuntat per les dades.

6.7.4 Mapa de relacions entre dades i objectes

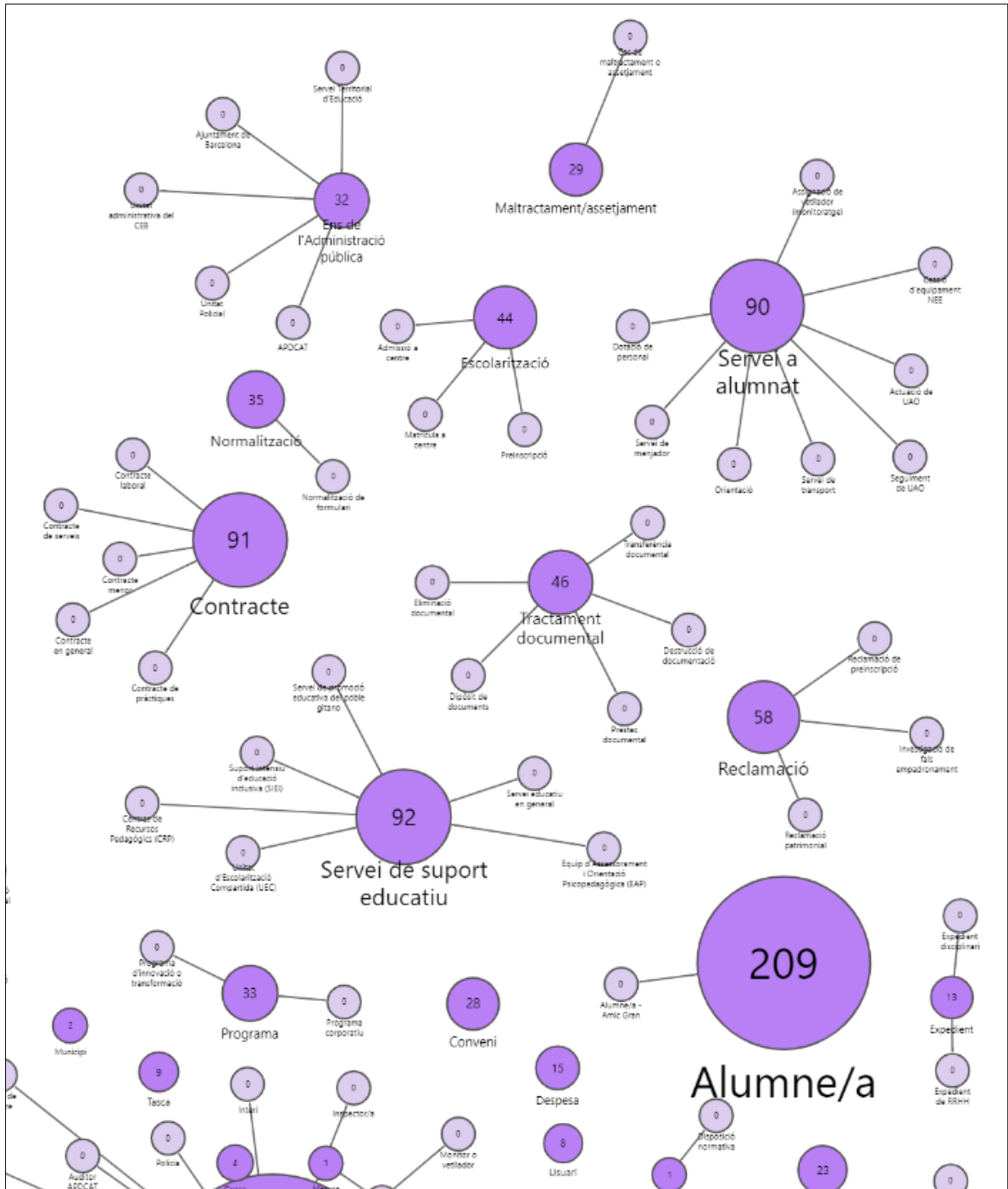
Per al Consorci s'ha elaborat tres representacions gràfiques. La primera relaciona els objectes amb les dades que els descriuen, per tal d'identificar visualment els objectes amb una major càrrega descriptiva.

Figura 6-8: Fragment del mapa de dades elementals del Consorci d'Educació de Barcelona



La segona representació agrupa els objectes en superobjectes, per tal de visualitzar les famílies d'objectes més rellevants, tant pel que fa a desplegament en objectes, com, a través de la dimensió, pel que fa al nombre de dades amb què es descriuen.

Figura 6-9: Fragment del mapa d'objectes del Consorci d'Educació de Barcelona



Capítol 7. Anàlisi dels resultats del cas d'aplicació

L'objectiu d'aquesta anàlisi és proporcionar evidències que permetin verificar o refutar les hipòtesis plantejades en el capítol 1 (Introducció). Aquestes evidències, que es recullen en el capítol 6, s'han obtingut a partir del cas d'aplicació del Consorci d'Educació de Barcelona, utilitzant per a la seva estructuració els conceptes i indicadors de la metodologia de governança de la informació descrita en els capítols 3, 4 i 5.

Malgrat haver demostrat la seva utilitat, les limitacions en la recollida d'informació, ja exposades en la introducció del capítol 6, no han permès en tots els casos obtenir una visió prou detallada de les relacions entre les entitats analitzades. Per tal de poder disposar d'aquesta visió, que permeti una anàlisi transversal del substrat informacional del Consorci, s'ha considerat convenient realitzar un estudi en profunditat sobre una mostra representativa de les sèries documentals del grup focal. Aquest estudi és el que s'exposa en l'apartat 7.1 d'aquest capítol.

La resta del capítol s'organitza d'acord amb els objectius teòrics fixats a la introducció d'aquesta tesi, per tal de facilitar posteriorment la seva utilització per a la verificació o refutació de les hipòtesis, i per a l'obtenció de conclusions generals.

7.1 Anàlisi de les característiques de les sèries documentals dadificades del Consorci d'Educació de Barcelona

Per seleccionar la mostra, s'ha considerat quatre indicadors de tria: els tres índex de dadificació (indicadors SER-13, SER-14, SER-15), i l'índex de ròssec intrasèrie (indicador SER-16). Les mètriques aplicades al grup focal són les següents:

Taula 7-1: Mètriques dels índex de dadificació i de ròssec de les sèries documentals del grup focal

SÈRIE DOCUMENTAL	ÍNDEX DE DADIFICACIÓ (EVIDÈNCIA)	ÍNDEX DE DADIFICACIÓ (DADES)	ÍNDEX DE DADIFICACIÓ (PONDERAT)	ÍNDEX DE RÒSSEC (INTRASÈRIE)
TRA004-001 Atenció a les peticions d'accés a la informació pública	0,38	0,68	0,51	0,70
TRA003-001 Assessorament en matèria de transparència als centres educatius	0,00	0,00	0,26	0,50

SÈRIE DOCUMENTAL	ÍNDEX DE DADIFICACIÓ (EVIDÈNCIA)	ÍNDEX DE DADIFICACIÓ (DADES)	ÍNDEX DE DADIFICACIÓ (PONDERAT)	ÍNDEX DE RÒSSEC (INTRASÈRIE)
TRA002-001 Gestió de les sol·licituds d'accés a documents i dades per part interessada	0,38	0,69	0,46	0,69
TRA001-001 Publicació d'informació a l'espai de transparència	1,00	1,00	0,56	0,00
ORG004-002 Normalització de formularis i models de document	0,00	0,00	0,00	0,00
ORG004-001 Normalització de formularis i models de documents	0,50	0,83	0,60	0,50
ORG003-001 Anàlisi i millora de processos	0,44	0,50	0,33	0,11
ORG002-001 Gestió de programes en l'àmbit de processos i sistemes	0,25	0,38	0,34	0,47
NEE021-001 Seguiment i valoració d'alumnes en situació de vulnerabilitat i/o malestar [...]	0,50	0,92	0,44	0,28
NEE020-001 Intervencions en matèria de seguretat ciutadana i prevenció de [...]	1,00	1,00	0,52	0,00
NEE002-001 Actuacions en casos de sospita de maltractament i/o assetjament en [...]	0,83	0,94	0,48	0,10
LOP007-001 Assessorament en matèria de protecció de dades personals	0,60	0,80	0,43	0,16
LOP006-001 Manteniment i publicació del registre d'activitats de tractament	0,33	0,10	0,27	0,47
LOP005-001 Auditories APDCAT i controls de seguretat sobre protecció de dades [...]	0,33	0,37	0,27	0,19
LOP003-001 Registre i notificacions de violacions de seguretat	0,33	0,28	0,17	0,08
LOP002-001 Atenció al drets d'autodeterminació informativa [...]	0,33	0,35	0,29	0,33
LOP001-001 Desenvolupament del sistema de protecció de dades de caràcter personal	0,40	0,23	0,20	0,15
INF009-002 Manteniment de les dades de la Base de dades de Centres (BDC)	0,00	0,00	0,56	1,00

SÈRIE DOCUMENTAL	ÍNDEX DE DADIFICACIÓ (EVIDÈNCIA)	ÍNDEX DE DADIFICACIÓ (DADES)	ÍNDEX DE DADIFICACIÓ (PONDERAT)	ÍNDEX DE RÒSSEC (INTRASÈRIE)
INF009-001 Manteniment de les dades de la Base de Dades de Centres (BDC)	0,60	0,95	0,46	0,00
INF007-001 Coordinació de l'ús social de centres	1,00	1,00	0,69	0,35
INF002-001 Eleccions al Consell Escolar de Centre	1,00	1,00	0,51	0,00
INF001-003 Elaboració d'estudis amb col·laboracions externes	0,25	0,65	0,28	0,18
INF001-002 Manteniment del banc de dades del CEB	0,80	0,81	0,42	0,02
INF001-001 Explotació de dades	0,63	0,89	0,59	0,50
GRH023-001 Gestió de borses de personal de suport als centres educatius (COVID-19)	1,00	1,00	0,79	0,57
GRH012-004 Comunicació d'exercici del dret de vaga	0,67	0,47	0,28	0,00
ESC001-001 Seguiment de falsos empadronaments	0,20	0,61	0,22	0,13
DOC004-001 Préstecs i consultes de documentació	0,44	0,88	0,65	0,64
DOC003-001 Transferències a l'Arxiu Central	0,33	0,73	0,64	0,72
DOC002-001 Tractament documental	0,25	0,53	0,44	0,67
DOC001-002 Assessorament als centres i serveis educatius en matèria de gestió [...]	0,80	0,92	0,77	0,70
DOC001-001 Desenvolupament del sistema de gestió documental	0,33	0,88	0,51	0,64

A continuació s'ha seleccionat d'aquest conjunt una mostra de deu sèries. Per garantir la representativitat d'aquestes sèries s'ha realitzat una anàlisi de conglomerats,²⁷¹ acotat a dos *clúster*, que ha donat el següent resultat:

Taula 7-2: Anàlisi clúster de les sèries documentals del grup focal

	ÍNDEX DE DADIFICACIÓ (EVIDÈNCIA)	ÍNDEX DE DADIFICACIÓ (DADES)	ÍNDEX DE DADIFICACIÓ (PONDERAT)	ÍNDEX DE RÒSSEC (INTRASÈRIE)
Clúster1	0,27	0,32	0,28	0,31
Clúster2	0,67	0,88	0,56	0,37

Sobre aquesta base s'ha triat les sèries de les quals es disposava d'informació de més detall, i que podien resultar més representatives de cada clúster. Aquestes sèries són les següents, per a les quals s'avança una atribució de caràcter i tipus de tramitació.²⁷²

Taula 7-3: Sèries documentals de la mostra extreta del grup focal

SÈRIE DOCUMENTAL	CLUSTER	CARÀCTER ESTIMAT	TIPUS DE TRAMITACIÓ
ESC001-001 Seguiment de falsos empadronaments	Clúster1	Transaccional	Reglada
GRH023-001 Gestió de borses de personal de suport als centres educatius (COVID-19)	Clúster2	Transaccional	Reglada
INF001-001 Explotació de dades	Clúster2	Transaccional	Reglada
INF009-001 Manteniment de les dades de la Base de Dades de Centres (BDC)	Clúster2	Recapitulativa	No reglada tècnica
LOP003-001 Registre i notificacions de violacions de seguretat	Clúster1	Transaccional	Reglada
LOP005-001 Auditories APDCAT i controls de seguretat sobre protecció de dades de caràcter personal	Clúster1	Transaccional	No reglada tècnica

²⁷¹ Per a les anàlisi de conglomerats s'utilitza en tots els casos l'algorisme *K-Means*.

²⁷² Informació desenvolupada en l'objectiu 5 (apartat 7.6 d'aquest capítol). Per a l'explicació dels conceptes de caràcter i tipus de tramitació de la sèrie documental, vegeu l'apartat 2.3.1.4.1 del capítol 2.

SÈRIE DOCUMENTAL	CLUSTER	CARÀCTER ESTIMAT	TIPUS DE TRAMITACIÓ
NEE002-001 Actuacions en casos de sospita de maltractament i/o assetjament en un centre educatiu	Clúster2	Transaccional	No reglada tècnica
NEE021-001 Seguiment i valoració d'alumnes en situació de vulnerabilitat i/o malestar social o psicosocial	Clúster1	Transaccional	Reglada
ORG003-001 Anàlisi i millora de processos	Clúster1	Transaccional	No reglada tècnica
ORG004-001 Normalització de formularis i models de documents	Clúster2	Transaccional	No reglada tècnica

Sobre aquesta mostra de deu sèries documentals s'ha portat a terme l'anàlisi per identificar les característiques d'una sèrie documental dadificada al Consorci d'Educació de Barcelona.

Als efectes d'aquesta anàlisi, per cada sèrie documental s'indica:

1. L'explicació dels factors de context que poden influir en les característiques de la sèrie, a través d'una breu descripció del procediment que genera la sèrie.
2. El detall del circuit evidencial de la sèrie, amb indicació de les evidències, tipus d'evidència, format i mitjà d'autenticació, i per aquelles que siguin del tipus dades indicació del sistema d'informació i la tecnologia.²⁷³
3. La matriu de variables de la sèrie, amb indicació de si l'ocurrència és amb caràcter de variable autòctona (AUT) o importada (IMP).
4. Els objectes que intervenen en la tramitació de la sèrie, en forma de *knowledge graph*. S'indiquen en color verd els objectes tipus "cosa", i en groc els objectes tipus "acció".
5. El mapa de relacions de la sèrie documental amb altres sèries, juntament amb el mapa de relacions del procés del qual depèn la sèrie. S'indiquen en color rosat els processos, i en verd les sèries documentals.

Cal tenir present que, amb la descripció de la sèrie, també s'estan descrivint les característiques de l'expedient tipus d'aquella sèrie. Passem doncs a descriure cadascuna de les sèries documentals de la mostra.

²⁷³ Per al detall tècnic dels formats, vegeu l'apartat 7.2 d'aquest capítol.

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

Per conèixer el motiu de l'ocultació vegeu l'apartat 1.6.2 del capítol 1.

7.1.9 Sèrie documental ORG003-001 Anàlisi i millora de processos

Qualsevol canvi en els processos de treball del Consorci implica la realització d'un projecte d'anàlisi i millora d'aquell procés. Aquest projecte s'inicia amb la recollida i descripció del procés, i posteriorment amb l'elaboració d'un informe de millores. En cas que es requereixi el desenvolupament d'una solució tecnològica, la unitat d'organització del Consorci elabora un informe de requeriments funcionals i el deriva als responsables del desenvolupament tecnològic. La descripció del procés s'informa al Catàleg de Processos del Consorci, i també al Catàleg Corporatiu de Processos de la Generalitat de Catalunya.

Taula 7-28: Circuit evidencial de la sèrie documental ORG003-001

EVIDÈNCIES	TIPUS	FORMAT	AUTENTIFICACIÓ
ORG003-E001 Sol·licitud d'anàlisi i millora de procés	Document		Metadades
ORG003-E002 Obertura de projecte d'anàlisi i millora de procés	Dades		Sense autenticació
ORG003-E003 Informe d'anàlisi i millora de procés	Document		Sense autenticació
ORG003-E004 Diagrama de procés	Document		Sense autenticació
ORG003-E005 Informe de requeriments funcionals de procés	Document		Sense autenticació
ORG003-E006 Tramesa d'informe d'anàlisi i millora de procés	Document		Metadades
ORG003-E007 Fitxa de procés al catàleg de processos	Dades		Sense autenticació
ORG003-E008 Fitxa de procés al catàleg corporatiu de processos	Dades		Metadades
ORG003-E009 Tancament de projecte d'anàlisi i millora de procés	Dades		Pista d'auditoria

Taula 7-29: Sistemes d'informació de la sèrie documental ORG003-001

EVIDÈNCIES	SISTEMES D'INFORMACIÓ	TECNOLOGIA
ORG003-E007 Fitxa de procés al catàleg de processos	SI02.012 Catàleg de processos del CEB	
ORG003-E008 Fitxa de procés al catàleg corporatiu de processos	SI05.007 Catàleg corporatiu de processos (OPAE)	
ORG003-E002 Obertura de projecte d'anàlisi i millora de procés	SI08.006 ORG003-001 Anàlisi i millora de processos	
ORG003-E009 Tancament de projecte d'anàlisi i millora de procés		

Taula 7-30: Matriu de variables de la sèrie documental ORG003-001

DADA ELEMENTAL	ORG00 3-E001	ORG00 3-E002	ORG00 3-E003	ORG00 3-E004	ORG00 3-E005	ORG00 3-E006	ORG00 3-E007	ORG00 3-E008	ORG00 3-E009
Nom del projecte	AUT	AUT				AUT	AUT	AUT	IMP
Nom i cognoms treballador del personal CEB	AUT		AUT		AUT	AUT			
Correu electrònic treballador CEB	AUT					AUT			
Nom unitat administrativa		AUT	AUT		AUT	AUT	AUT	AUT	IMP
Codi del projecte		AUT					AUT	AUT	IMP
Data d'inici del projecte		AUT					AUT		IMP
Estat del projecte		AUT							AUT
Nom de l'informe			AUT						
Proposta de procediment			AUT						
Fluxgrama de procés				AUT					
Canvis realitzats					AUT				
Data de l'informe					AUT				
Versió					AUT				
Data de tancament del projecte							AUT		AUT
Tipus de procés							AUT		
Descripció de procés								AUT	

Figura 7-17: Objectes descrits en la sèrie documental ORG003-001

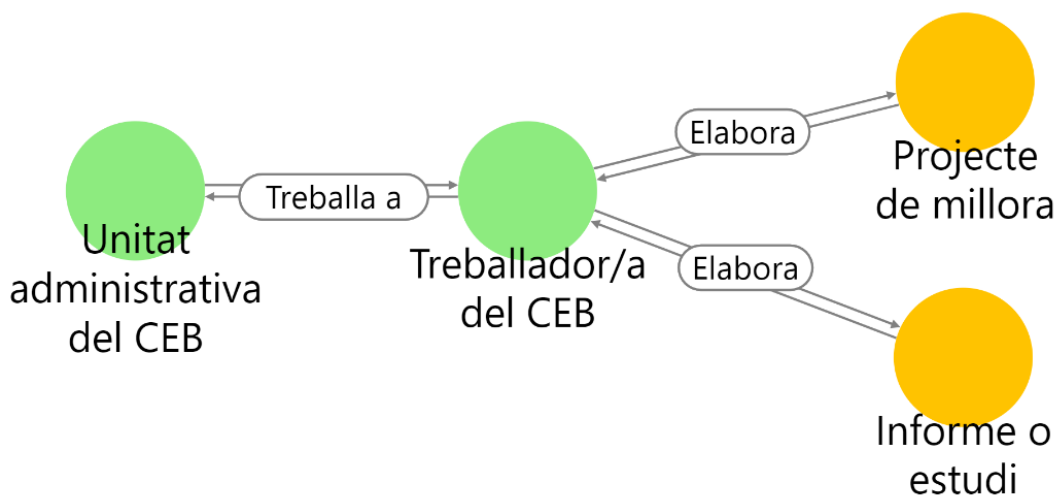
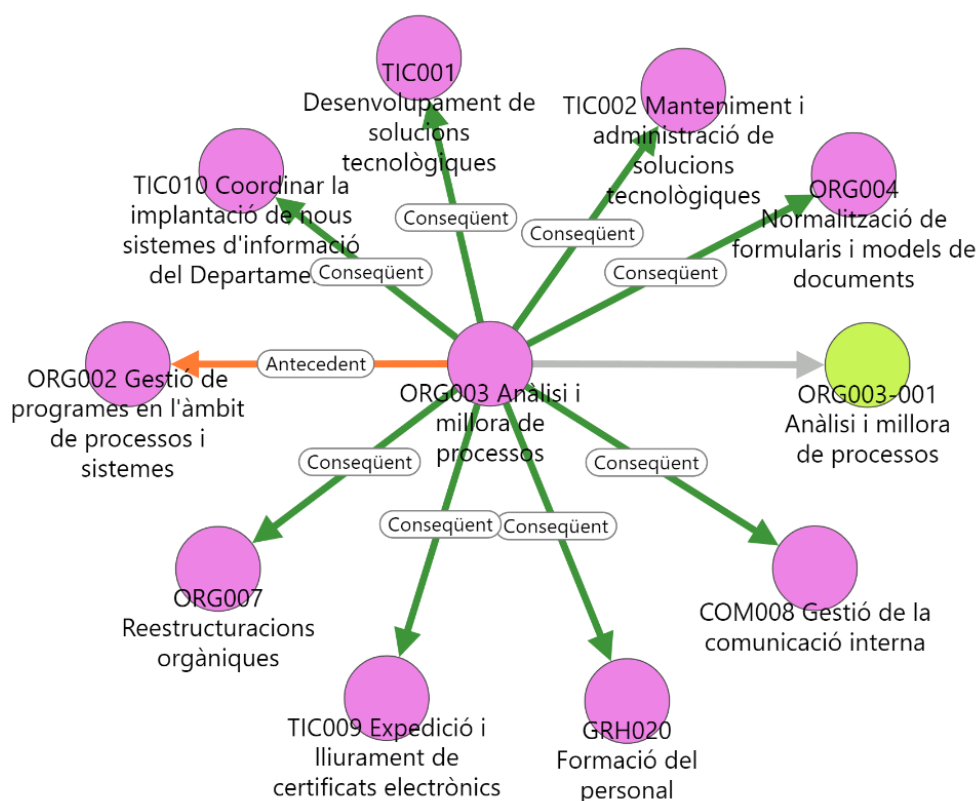


Figura 7-18: Mapa de relacions del procés i la sèrie documental ORG003-001



7.1.10 Sèrie documental ORG004-001 Normalització de formularis i models de documents

Qualsevol formulari o model de document que s'hagi d'utilitzar al Consorci passa per un procés de normalització, prèvia aprovació per part de la Direcció de la seva necessitat. Un cop elaborat el formulari o model de document, es valida amb la unitat que en serà usuària, s'implementa tecnològicament, i es publica en el canal pertinent, en funció de si va adreçat a ciutadania, a centres educatius, o si és d'ús intern. El formulari es registra en un registre propi del Consorci.

Taula 7-31: Circuit evidencial de la sèrie documental ORG004-001

EVIDÈNCIES ²⁷⁵	TIPUS	FORMAT	AUTENTIFICACIÓ
ORG004-E001 Sol·licitud de normalització de formulari o model de document	Dades	[REDACTED]	Metadades
ORG004-E002 Registre sol·licitud normalització de formulari o model de document	Dades	[REDACTED]	Sense autenticació
ORG004-E003 Comunicació recepció sol·licitud de normalització	Document	[REDACTED]	Metadades
ORG004-E004 Aprovació/ denegació de sol·licitud de normalització	Dades	[REDACTED]	Pista d'auditoria
ORG004-E005 Registre del desenvolupament del formulari o model de document	Dades	[REDACTED]	Pista d'auditoria
ORG004-E006 Formulari o model de document	Document	[REDACTED]	Sense autenticació
ORG004-E007 Sol·licitud de validació de formulari o model de document	Document	[REDACTED]	Metadades
ORG004-E008 Validació de formulari o model de document	Document	[REDACTED]	Metadades
ORG004-E010 Petició publicació formulari o model de document	Document	[REDACTED]	Metadades
ORG004-E011 Comunicació publicació formulació o model de document	Document	[REDACTED]	Metadades
ORG004-E012 Registre de la publicació del formulació o model de document	Dades	[REDACTED]	Sense autenticació

Taula 7-32: Sistemes d'informació de la sèrie documental ORG004-001

EVIDÈNCIES	SISTEMES D'INFORMACIÓ	TECNOLOGIA
ORG004-E001 Sol·licitud de normalització de formulari o model de document	SI08.007 ORG004-001 Normalització de formularis i models de documents	[REDACTED]
ORG004-E002 Registre sol·licitud normalització de formulari o model de document		
ORG004-E004 Aprovació/ denegació de sol·licitud de normalització		
ORG004-E005 Registre del desenvolupament del formulari o model de document		
ORG004-E012 Registre de la publicació del formulació o model de document		

Taula 7-33: Matriu de variables de la sèrie documental ORG004-001

DADA ELEMENTAL	ORG0 04-E001	ORG0 04-E002	ORG0 04-E003	ORG0 04-E004	ORG0 04-E005	ORG0 04-E006	ORG0 04-E007	ORG0 04-E008	ORG0 04-E010	ORG0 04-E011	ORG0 04-E012
Nom i cognoms treballador CEB	AUT	IMP	IMP	IMP	IMP		AUT		AUT	AUT	IMP
Nom unitat administrativa	AUT	IMP	IMP	IMP	IMP		AUT		AUT	AUT	IMP
Ús previst	AUT	IMP	IMP	IMP	IMP						IMP
Títol formulari	AUT	IMP		IMP	IMP					AUT	IMP

²⁷⁵ El salt en la numeració de les evidències obeeix a una evidència que ha deixat de ser vigent un cop finalitzada la transformació digital de la sèrie documental.

DADA ELEMENTAL	ORG0 04-E001	ORG0 04-E002	ORG0 04-E003	ORG0 04-E004	ORG0 04-E005	ORG0 04-E006	ORG0 04-E007	ORG0 04-E008	ORG0 04-E010	ORG0 04-E011	ORG0 04-E012
Data lliurament	AUT	IMP		IMP	IMP						IMP
Exemple de formulari	AUT	IMP								AUT	
Correu electrònic treballador CEB	AUT		AUT	IMP	IMP						IMP
Descripció formulari	AUT										
Funcionalitats	AUT										
Data encàrrec		AUT	IMP	IMP	IMP						IMP
Estat de petició		AUT		AUT	IMP					AUT	IMP
Estat tasca - (ORG004)		AUT		AUT	IMP					AUT	IMP
Codi formulari		AUT		IMP	IMP					AUT	IMP
Format formulari		AUT		IMP	IMP					AUT	IMP
Nom del procés		AUT		IMP	IMP					AUT	IMP
Observacions (ORG004)		AUT		IMP	IMP					AUT	IMP
Prioritat tasca		AUT		IMP	IMP					AUT	IMP
Descripcions de la sol·licitud		IMP	AUT	IMP	IMP						IMP
Descripció funcionalitat		IMP									
Número de petició			AUT	IMP	IMP						IMP
Prioritat del lliurament			AUT								
Tipus de demanda			AUT								
Aprovat/denegat				AUT	IMP		AUT	AUT			IMP
Data desenvolupament					AUT						IMP
Tipus document					AUT						IMP
Nom formulari						AUT	AUT	AUT	AUT		
Id. formulari - formulari document						AUT					
Sol·licitud de publicació									AUT		
Data publicació										AUT	AUT
Publicació de formulari										AUT	

Figura 7-19: Objectes descrits en la sèrie documental ORG004-001

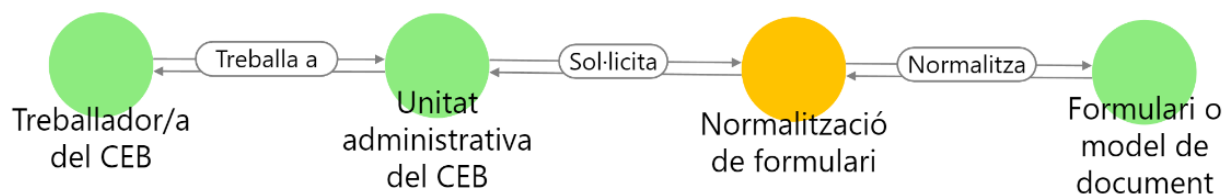
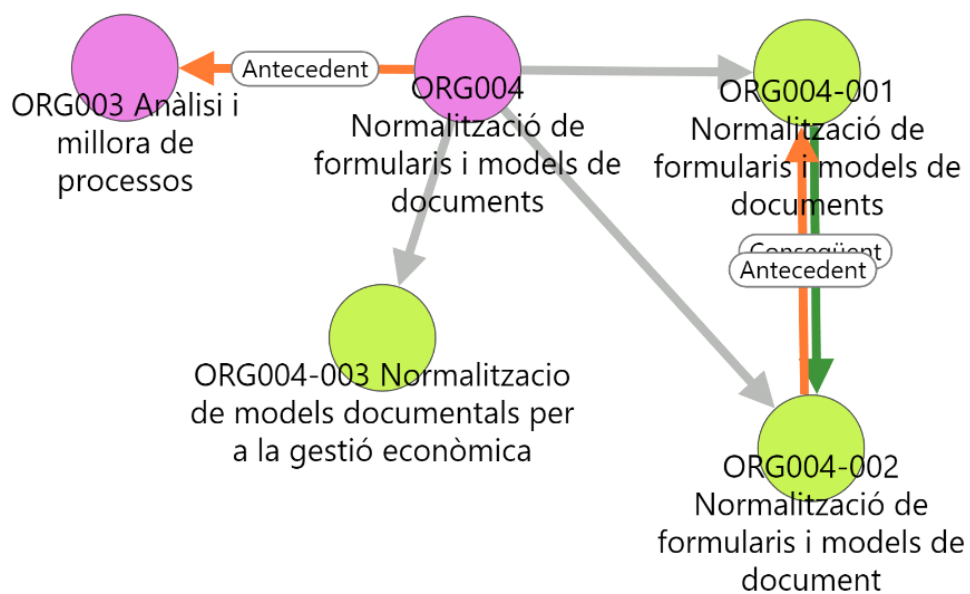


Figura 7-20: Mapa de relacions del procés i la sèrie documental ORG004-001



7.2 Objectiu 1: característiques de la sèrie documental dadificada

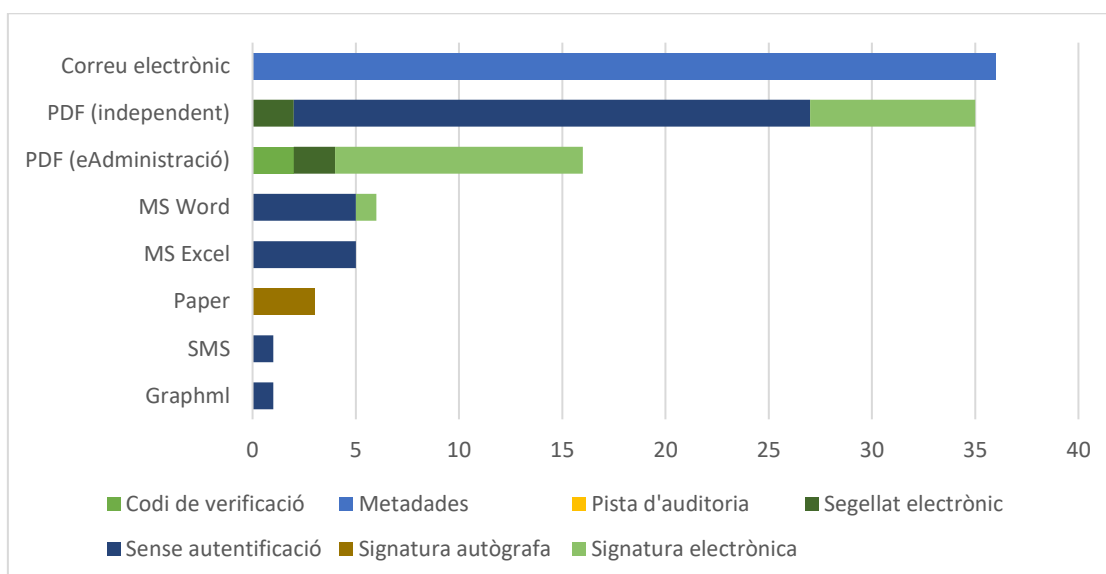
El primer objectiu consisteix en identificar si existeix un patró en les característiques formals i substantives de les sèries documentals electròniques que han passat per un procés de dadificació. Considerem una sèrie documental com electrònica quan totes les evidències que conserva l'organització d'aquella sèrie estan en format electrònic.²⁷⁶ I la considerarem com a sèrie completament dadificada quan totes les seves evidències estiguin en format de dades estructurades i emmagatzemades en sistemes d'informació. Cal tenir present que, mentre que el caràcter digital de la sèrie és un atribut absolut (és digital o no), el caràcter de sèrie dadificada és un atribut relatiu, atès que no sempre és possible, o bé òptima, la dadificació completa d'una sèrie documental.

Per poder obtenir aquest patró a partir de les dades del Consorci hem combinat dues fonts d'informació. En primer lloc, les dades completes de les 32 sèries del grup focal, atès que són les que disposen de la descripció completa informada a l'eina de recollida, i a més han completat en tots els casos un procés de dadificació, completa o parcial segons el cas. I en segon lloc les dades detallades de les 10 sèries documentals de la mostra que ha estat objecte d'estudi específic. I també utilitzarem com a font el mapa de sistemes, el mapa de sèries i el mapa de dades del Consorci, atès que disposen cadascun de la informació completa en el seu àmbit d'actuació.

Si analitzem les característiques formals de les evidències generades com a evidències-objecte per les sèries del grup focal, observem el següent:

²⁷⁶ Es poden generar còpies de les evidències de la sèrie en format paper o en altres formats si les necessitats o preferències d'agents externs ho requereixen, com s'esdevé en el sector públic arran de les obligacions de compliment de l'article 14.1 de la Ley 39/2015, de 1 de octubre, del Procedimiento Administrativo Común de las Administraciones Públicas, el qual estableix que "Las personas físicas podrán elegir en todo momento si se comunican con las Administraciones Públicas para el ejercicio de sus derechos y obligaciones a través de medios electrónicos o no [...]". Mentre els exemplars no digitals estiguin custodiats per agents diferents al productor de la sèrie, aquest fet no afectarà el caràcter digital de la sèrie documental.

Gràfic 7-1: Formats i mitjans d'autenticació de les evidències-objecte del grup focal



Malgrat que un 54% de les evidències del grup focal (un conjunt de sèries dadificades) són evidències-objecte, el 71% d'aquestes evidències corresponen a missatgeria (correus electrònics i SMS) o bé a documents que no tenen cap mitjà d'autenticació. Es tracta doncs de documents que comuniquen dades, però no les consoliden en cap forma fefaent.

Només els documents PDF que es trameten a través d'algun canal d'Administració electrònica disposen en tots els casos d'algun mitjà d'autenticació reconegut legalment.²⁷⁷ L'ús de la signatura electrònica basada en certificat, com a sistema d'autenticació forta, es concentra doncs en aquests PDF tramesos per canals d'eAdministració, i en una tercera part dels PDF interns, en la majoria dels casos informes de constància interna. En la resta de casos l'ús de mitjans d'autenticació és residual o inexistent.

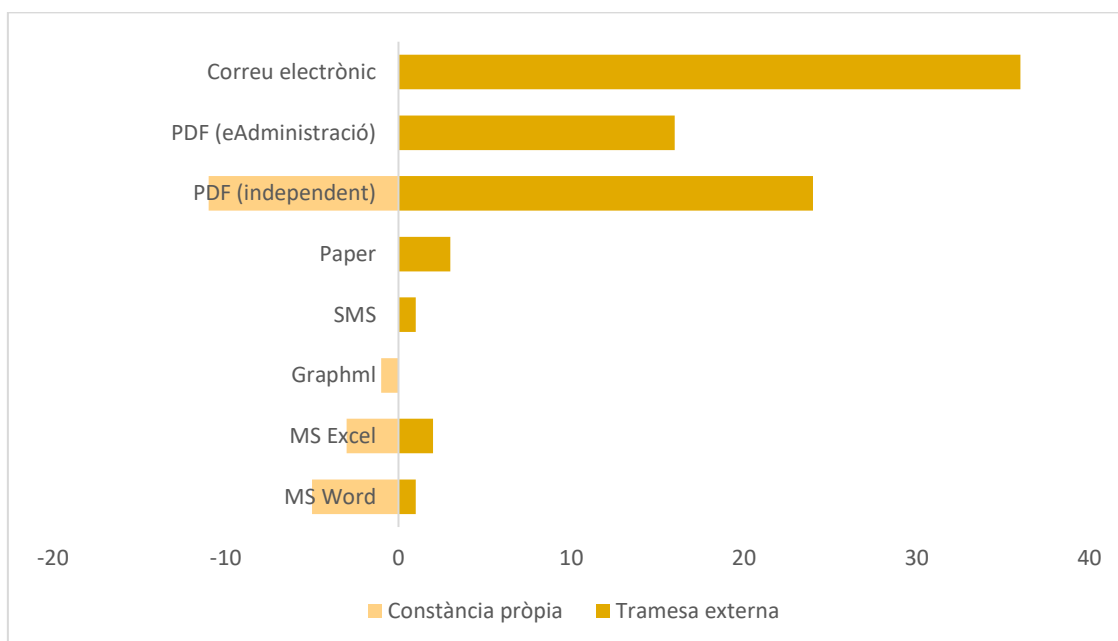
Si ens fixem en els usos de les evidències-objecte, que podem detectar a través de la cadena productor-receptor del document, observem que la majoria de les evidències han estat creades per a la tramesa externa (recepció o emissió), a partir de dades de les quals el Consorci ja disposa en format estructurat. Només en uns pocs casos les evidències tenen el mateix productor i receptor, és a dir actuen com a elements de

²⁷⁷ Els mitjans contemplats són els serveis eValisa, eNotum, EACAT i GSIT, tots ells desenvolupats per la Generalitat de Catalunya i del Consorci de l'Administració Oberta de Catalunya. eValisa és un sistema de comunicació intern de les administracions (<https://www.aoc.cat/serveis-aoc/e-valisa/>), eNotum és la plataforma de notificació electrònica (<https://www.aoc.cat/serveis-aoc/e-notum/>), EACAT és l'extranet de comunicació entre les administracions catalanes (<https://www.aoc.cat/serveis-aoc/eacat-tramits/>), i GSIT (Gencat Serveis i Tràmits) és la solució per a la interacció amb ciutadania (<https://ctti.gencat.cat/ca/detalls/detallarticle/gsit>).

Pel que fa als formats, en el Consorci s'ha identificat fitxers en format Acrobat PDF 1.6 (PUID fmt/20), Acrobat PDF 1.7 (PUID fmt/276), Microsoft Word en OOXML (PUID fmt/1827) i Microsoft Excel en OOXML (PUID fmt/1828).

constància dins l'expedient. Aquests casos corresponen a fixacions en PDF d'informes d'ús exclusivament intern, o a documents en format ofimàtic subjectes a reutilització.

Gràfic 7-2: Finalitat en la creació de les evidències-objecte del grup focal



Arran d'aquesta observació podem convenir que, malgrat la pervivència d'evidències-objecte en la sèrie documental dadificada en un percentatge superior al 50% de les evidències, el pes provatori ja s'ha traslladat a la informació estructurada, que rau en sistemes d'informació amb un nivell de seguretat suficient, i en tots els casos amb mecanismes d'autenticació.

Taula 7-34: Nivell de seguretat dels sistemes d'informació del grup focal

ABAST	SISTEMES	SEGURETAT (SIS-20)
Usuari final	1	1
Ajuntament de Barcelona	1	1
CEB empresa externa	1	1
CEB propi	29	1
Departament d'Ensenyament	1	1

Hem cregut necessari completar aquesta informació amb una comparativa de dues de les sèries de la mostra amb nivells més dispars de dadificació, per veure què tenen en comú i què les diferencia. Hem triat com a més dadificada la sèrie documental GRH023-001 Gestió de borses de personal de suport als centres educatius (COVID-19) (índex

SER-15=0,79), i com a menys dadificada la sèrie documental LOP003-001 Registre i notificacions de violacions de seguretat (índex SER-15=0,17). Estructurarem la comparativa pels diferents ítems de la descripció de cada sèrie: evidències, sistemes d'informació, objectes i agents externs.

El **nombre d'evidències**, que ve determinat per la complexitat del procediment, i per la quantitat de tasques o accions que es realitzen en el flux de treball, coincideix entre ambdues sèries: GRH023-001 en té 6 (totes de dades), i LOP003-001 en té també 6 (4 evidències-objecte i 2 de dades). La dadificació del substrat evidencial no altera en principi aquesta mètrica, atès que no actua sobre la naturalesa del procediment, sinó només sobre el seu registre evidencial.

També coincideixen ambdues sèries en el fet de tenir com a base de les evidències dadificades un únic **sistema d'informació** basat en un desenvolupament *low code*. L'ús d'aquest sistema de registre no altera tampoc la forma de la matriu de variables, llevat de la diferència en la intensitat de la reutilització de les dades (ròssec intrasèrie).

Taula 7-35: Matriu de variables de la sèrie documental GRH023-001

	GRH023- E001	GRH023- E002	GRH023- E003	GRH023- E004	GRH023- E005	GRH023- E006
Codi de centre	AUT		IMP	IMP	IMP	IMP
Nom de centre	AUT		IMP	IMP	IMP	IMP
Districte de centre	AUT		IMP			
Correu electrònic de centre	AUT			IMP	IMP	
Telèfon de centre	AUT			IMP	IMP	
Nom i cognoms director centre del docent	AUT			IMP		
Barri del centre	AUT					
Telèfon del director	AUT					
Titularitat de centre	AUT					
DNI/NIF del Docent		AUT		IMP	IMP	IMP
Nom i cognoms del docent		AUT		IMP	IMP	IMP
Correu electrònic del docent		AUT		IMP	IMP	
Adreça particular		AUT				
Data naixement del Docent		AUT				
Telèfon corporatiu del docent		AUT				
Data petició substitució			AUT	IMP	IMP	
Especialitat del Docent			AUT	IMP	IMP	
Motiu substitució			AUT	IMP	IMP	
Motiu de trucada			AUT			
Número de substitució				AUT	IMP	IMP
Data fi de substitució				AUT	IMP	
Justificació de la petició				AUT	IMP	
Valoració de substitució					AUT	IMP
Observacions centre					AUT	
Alerta						AUT

Taula 7-36: Matriu de variables de la sèrie documental LOP003-001

	LOP003-E001	LOP003-E002	LOP003-E003	LOP003-E004	LOP003-E005	LOP003-E006
Nom i cognoms treballador del personal CEB	AUT	IMP	IMP	IMP		IMP
Nom unitat administrativa	AUT	IMP		IMP	IMP	IMP
Data de l'esdeveniment	AUT		IMP	IMP	IMP	IMP
Mesures aplicades	AUT			IMP		IMP
Canal de comunicació	AUT			IMP		
Tipus d'esdeveniments	AUT					IMP
Codi violació de seguretat		AUT				IMP
Data de comunicació			AUT	IMP		IMP
ID denuncia			AUT		IMP	
Circumstàncies del fet			AUT			
Data de naixement del treballador del personal CEB			AUT			
DNI/NIF treballador del personal CEB			AUT			
Llista d'objectes sostrets			AUT			
Lloc de l'esdeveniment			AUT			
Lloc de naixement treballador/a CEB			AUT			
Nom i cognoms treballador/a policia			AUT			
Nom unitat policial			AUT			
Descripció de l'incident				AUT	IMP	
Adreça de l'entitat				AUT		
Càrrec del treballador del personal CEB				AUT		
Categoria de dades personals afectades				AUT		
CIF d'entitat				AUT		
Codi postal de l'entitat				AUT		
Col·lectiu de les persones afectades				AUT		
Comunicació de la VS a altres organismes				AUT		
Contingut de la comunicació				AUT		
Correu electrònic treballador del personal CEB				AUT		
Naturalesa de la VS				AUT		
Nombre de persones afectades				AUT		
Nombre de persones afectades que es preveu trametre la comunicació				AUT		
Nombres de registres de dades personals afectades				AUT		
Població de l'entitat				AUT		
Possible conseqüències de la VS				AUT		
S'ha comunicat				AUT		
Telèfon fix del personal del CEB				AUT		
Tipus de notificació				AUT		
Tipus d'entitat				AUT		
Forma de comunicació					AUT	IMP
Nom administració APDCAT					AUT	
Tipus de document incidència seguretat						AUT

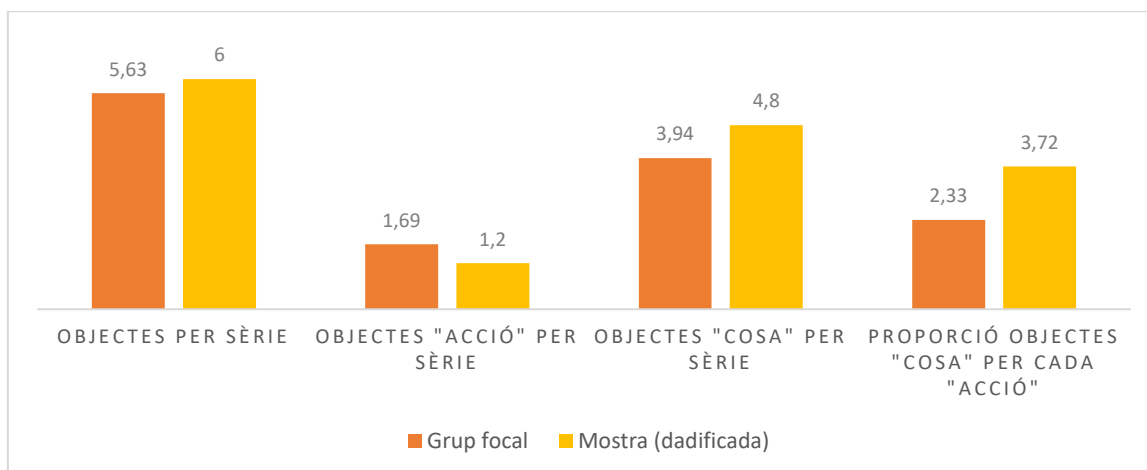
No s'observen tampoc diferències significatives en el nombre d'objectes descrits (GRH023-001 descriu 4 objectes, 3 coses i 1 acció, mentre que LOP003-001 en descriu 6, 5 coses i 1 acció), ni en el vincle amb els processos (en ambdós casos existeix relació amb un procés sense relacions complexes).

La principal diferència es troba en la **naturalesa de les evidències**. El principal determinant en aquest cas és l'existència en el procediment d'agents externs amb els quals calgui intercanviar informació. Si aquests agents són indeterminats, no és possible gestionar un accés a un sistema d'informació comú, o en tot cas requereix nivell d'autenticació més alts, i per tant una major complexitat tecnològica, que una solució *low code* sovint no pot cobrir. Per aquest motiu, la sèrie LOP003-001, en la qual, a diferència de la sèrie GRH023-001, participen agents externs (APDCAT, unitat policial, etc.), requereix la generació d'evidències objecte amb autenticació pròpia. En aquest sentit existeixen també factors culturals que analitzem més endavant.

Això repercuteix també en el nivell d'integració entre les evidències, i per tant de reutilització de les dades (el ròssec intrasèrie), que baixa en funció de la quantitat d'agents externs no determinables. Veiem doncs que la principal diferència, i el que condiciona directament la dadificació dels processos, encarint i fent més complexa la dadificació, és la possibilitat d'obrir els sistemes d'informació a agents externs a l'organització. Quan tots els agents pertanyen a la pròpia organització i es poden autenticar amb facilitat, la dadificació completa és un objectiu més assolible.

Una darrera característica que volem analitzar és la relació entre objectes de tipus "cosa" i objectes de tipus "acció", comparant en aquest cas les mesures del grup focal (32 sèries), ja analitzades amb l'indicador SER-08 (apartat 6.4.3 del capítol 6) amb les mesures de la mostra (10 sèries). Cal tenir present que la mostra inclou sèries dadificades, algunes intensivament, on s'ha detallat al màxim la informació sobre les dades gestionades. Així doncs, si comparem les mitjanes del grup focal amb les de la mostra, observem el següent:

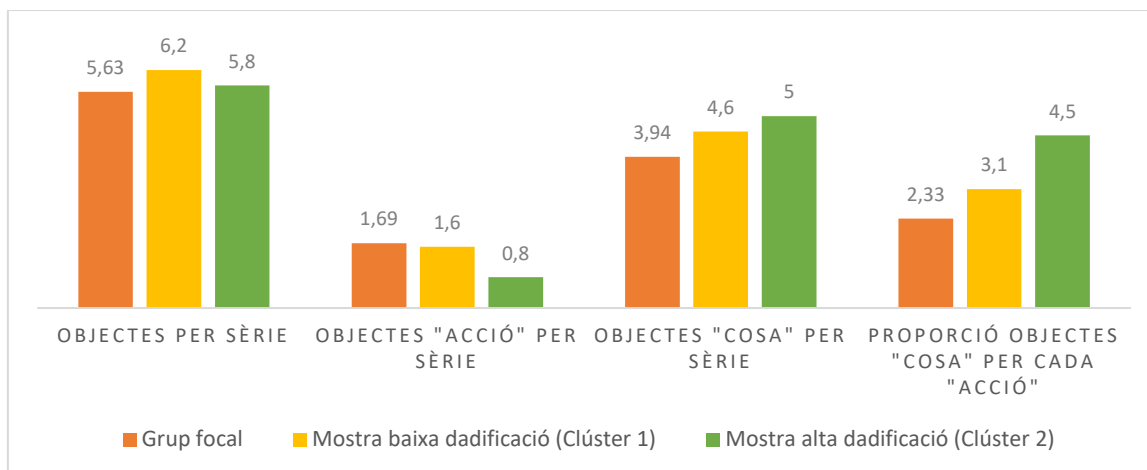
Gràfic 7-3: Comparativa de les mètriques d'objectes del grup focal i de la mostra



Veiem que, en general, les sèries documentals de la mostra superen la mitjana del grup focal en el nombre d'objectes tipus "cosa", com reflecteix la proporció de "coses" descrites per cada acció, que supera en més d'un punt la mitjana del grup focal.

Però cal tenir en compte que la mostra agrupa dos tipus de sèries: les sèries del clúster 2, on s'ha pogut assolir un alt nivell de dadificació, i les sèries del clúster 1, on els nivells de dadificació que s'han pogut assolir són notablement més baixos. Incorporant aquest factor de correcció s'obtenen les següents dades:

Gràfic 7-4: Comparativa de les mètriques d'objectes del grup focal i de la mostra diferenciades per clúster



Això ens permet observar més clarament els efectes que una dadificació intensiva ha provocat sobre les sèries documentals del Consorci d'Educació de Barcelona, que podem condensar en:

1. Una racionalització del nombre d'objectes per sèrie, ajustada a la mitjana.
2. Un ajustament a l'òptim funcional, segons el qual cada sèrie documental fa referència a una única activitat, i a cap si es tracta d'una sèrie recapitulativa.²⁷⁸
3. Un increment del nombre de "coses" sobre les quals la sèrie documental recull informació, el que denota probablement una major riquesa de dades contextuais per emmarcar l'acció.

Aquest increment d'informació sobre coses, i per tant de dades disponibles per a prendre decisions sobre l'acció de la sèrie documental, també és objecte d'anàlisi en l'objectiu 6 arran de la introducció de les visualitzacions de dades com nous components de la sèrie documental.

²⁷⁸ L'indicador de 0,8 (inferior a 1) el proporciona la sèrie documental INF009-001, una sèrie clarament recapitulativa que no recull dades de cap objecte "acció".

7.3 Objectiu 2: característiques de l'ecosistema tecnològic de la sèrie documental dadificada

El segon objectiu és l'anàlisi de l'ecosistema tecnològic que dona suport a les sèries documentals dadificades. Si analitzem les solucions utilitzades per les sèries del grup focal del Consorci, observarem com un 64% es basa en tecnologia de col·laboració

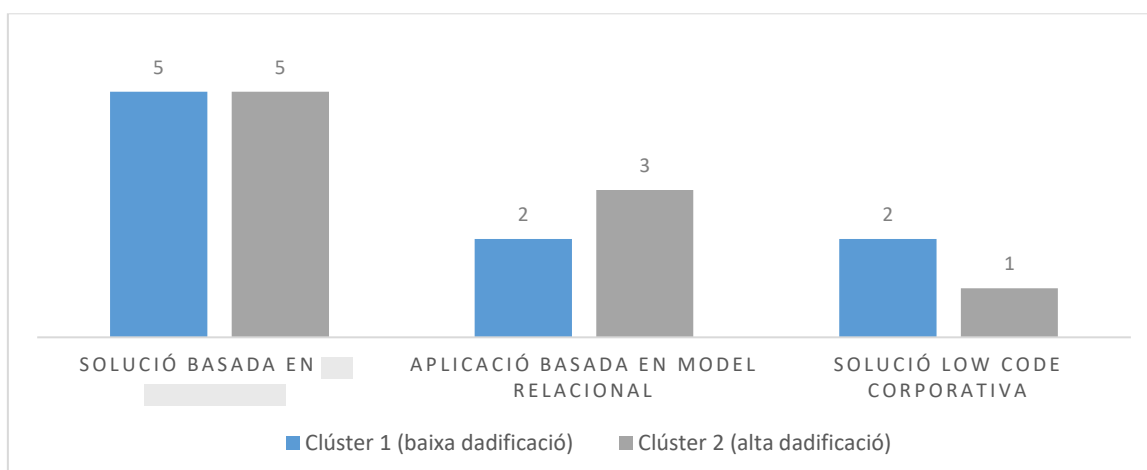
Ens aturem un moment a desgranar els components que s'han desenvolupat en algunes d'aquestes solucions:

- [Redacted]
- [Redacted]
- [Redacted]
- [Redacted]

Tenint present que el grup focal concentra sèries sobre les quals s'ha pogut actuar en transformació digital, especialment en els tres darrers anys, podem deduir que l'aposta que ha fet el Consorci d'Educació de Barcelona pel desenvolupament *low code*, [Redacted] ha permès accelerar la transformació digital d'una forma important.

Si traslладem l'anàlisi a la mostra, no observem diferències significatives pel que fa a la tecnologia utilitzada entre les sèries amb diferents nivells de dadificació (clúster 1 i clúster 2):

Gràfic 7-5: Comparativa del nombre de sistemes d'informació per tecnologia i clúster



Convenim doncs que l'escenari tecnològic més habitual de la sèrie documental dadificada, amb independència del seu nivell de dadificació, inclou:

- En tots els casos, un desenvolupament *low code* en un entorn de col·laboració, probablement per absorbir i ordenar el que fins ara es gestionava amb ofimàtica (vegeu taula 6-44 del capítol 6). Aquests desenvolupaments són majoritàriament desenvolupats internament en modalitat *Software as a Service* (vegeu taula 6-39 i 6-40 del capítol 6).
- Aquesta tria tecnològica també fa que l'entorn d'emmagatzemament que estigui creixent més són les base de dades orientades a objectes (vegeu taula 6-46 i gràfic 6-19 del capítol 6).
- Accés quan és necessari a aplicacions consolidades, sovint amb caràcter de dades mestres. Sovint aquest accés no es pot fer directament, i requereix l'ús de robòtica administrativa (vegeu taula 6-52 del capítol 6).
- Utilització de solucions *low code* corporatives quan són necessàries interaccions amb agents externs (vegeu les característiques de la sèrie documental dadificada, a l'apartat 7.2 d'aquest capítol).

7.4 Objectiu 3: equivalència entre la sèrie documental i el sistema d'informació

El tercer objectiu requereix analitzar si existeix una correlació entre el concepte de sèrie documental i el de sistema d'informació. La hipòtesi inicial és que els sistemes d'informació d'una organització són equiparables a agrupacions documentals, i específicament al concepte de sèrie documental (és a dir, es compleix l'equiparació sèrie-sistema), i per tant poden ser objecte, amb les degudes adaptacions, del tractament propi de la gestió documental (conservació/eliminació i accés)

En el cas del Consorci s'ha fet aquesta anàlisi des de dues dimensions:

1. La possible correspondència 1 a 1 entre sèrie documental i sistema d'informació.
2. El nivell de coincidència entre les matrius de dades de la sèrie i del sistema.

En el primer punt, s'infereix que el fet que un sistema d'informació doni suport només a una sèrie documental és un possible indicador que el sistema pot ser assimilable a sèrie documental. I aquest el complementem amb el segon, incorporant l'índex de coincidència entre les matrius o models de dades de la sèrie documental i del sistema d'informació, tot inferint que un alt índex de coincidència faria possible que el sistema d'informació actués directament com a sèrie documental, fent innecessària la generació d'evidències-objecte (llevat de les imprescindibles per a la comunicació externa).

L'indicador SIS-06²⁷⁹ ja apunta aquesta tendència: el 83% dels sistemes d'informació relacionats amb sèries ho són només amb una única sèrie. I del conjunt de sistemes vinculats a sèries, el 62% correspon a desenvolupaments *low code*, i el 24% a ofimàtica.

Analitzem a continuació cadascun dels dos punts, agrupant l'anàlisi per tecnologia d'emmagatzemament.²⁸⁰

7.4.1 Equivalència amb el model relacional

Entenem per model relacional aquelles aplicacions, de gestió o registrals, que utilitzen majoritàriament bases de dades relacionals per emmagatzemar la seva informació.

Els sistemes d'informació basats en tecnologia relacional tenen una ràtio de 1,47 sèries per sistema, mentre que les sèries utilitzen de mitjana 1,17 sistemes per sèrie.

Només en un cas s'ha trobat coincidència de dades entre la sèrie documental i el sistema d'informació (INF009-001 Manteniment de les dades de la Base de Dades de Centres

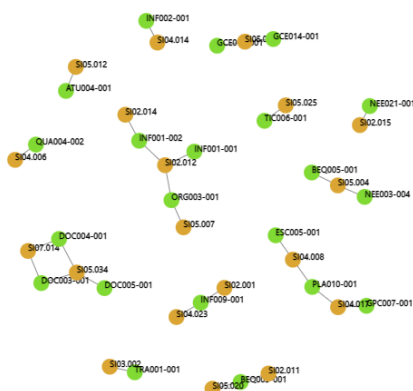
²⁷⁹ Vegeu gràfics 6-18 i 6-19 del capítol 6.

²⁸⁰ En la representació gràfica es representen en groc els sistemes i en verd les sèries.

(BDC) amb SI02.001 Base de Dades de Centres (BDC)), amb un índex de coincidència força elevat: 0,63. Atès que es tracta d'una sèrie que podem considerar recapitulativa, analitzem aquest cas en l'objectiu 4 (apartat 7.5 d'aquest capítol).

La representació gràfica de les relacions en les sèries del grup focal és la següent:

Figura 7-21: Mapa de relacions de sèries documentals amb sistemes d'informació del model relacional



7.4.2 Equivalència amb el model *data warehouse*

Entenem per model *data warehouse* aquells conjunts de dades de gran volum, els quals a través de diferents tècniques de modelat es disposen en una estructura orientada a l'explotació (*datamarts*), utilitzant sistemes de *business intelligence* (BI).

No s'ha pogut identificar cap relació dels sistemes tipus *Data warehouse* amb sèries documentals, atès que no s'utilitzen en el marc de cap procés. No podem valorar doncs la possible equivalència.

7.4.3 Equivalència amb el model *data lake*

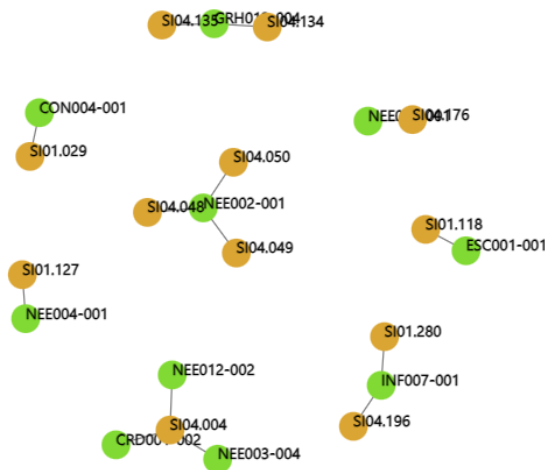
Entenem per model *data lake* aquells conjunts de dades estructurades, sense una estructura o modelat complex, que s'emmagatzemen i es consumeixen en entorns tipus llac de dades. Habitualment es tracten com taules de dades simples, i en tot cas en aquest punt limitarem l'anàlisi als conjunts de dades d'origen preferentment ofimàtic.

Els sistemes d'informació basats en conjunts de dades simples es relacionen amb una mitjana de 1,16 sèries per sistema. Cada sèrie documental que es basa en aquest tipus de sistemes utilitza 1,4 sistemes per sèrie.

En 8 casos s'ha trobat coincidència de dades entre conjunts de dades simples i sèries documentals, amb un índex de coincidència de 0,44.

La representació gràfica de les relacions en les sèries del grup focal és la següent:

Figura 7-22: Mapa de relacions de sèries documentals amb sistemes d'informació del model distribuït



7.4.4 Equivalència amb el model orientat a objectes

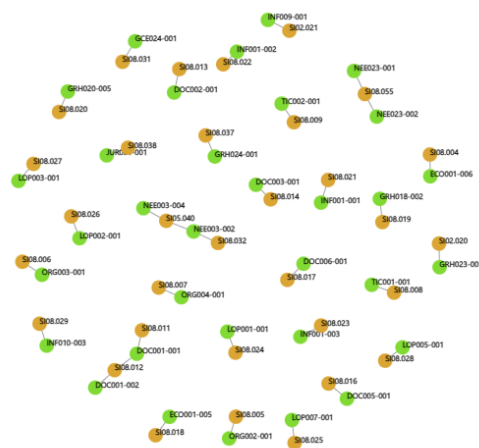
Entenem per model orientat a objectes aquells conjunts d'objectes metadescrits, que poden incloure contingut (fitxers) o únicament metadades, i que es gestionen en plataformes amb funcionalitats de gestor documental (EDMS). Aquest model inclou els entorns de col·laboració i compartició de continguts .

La correspondència més alta entre sistema i sèrie documental es dona en aquesta tecnologia. Els sistemes tipus gestor documental s'associen a una mitjana de 1,09 sèries per sistema, mentre que les sèries utilitzen de mitjana 1,06 sistemes per sèrie.

En 20 casos s'ha trobat coincidència entre les dades de la sèrie documental i les del sistema d'informació, però amb un índex mitjà relativament baix, del 0,33.

La representació gràfica de les relacions en les sèries del grup focal és la següent:

Figura 7-23: Mapa de relacions de sèries documentals amb sistemes d'informació del model orientat a objectes



7.4.5 Conclusió

La síntesi de les dades obtingudes és la següent:

Taula 7-37: Índex de coincidència sèrie documental - sistema d'informació per model tecnològic

	Sèries per sistema	Sistemes per sèrie	Casos de coincidència	Índex de coincidència
Model relacional	1,47	1,17	1 de 13	0,63
Model <i>data lake</i>	1,16	1,40	8 de 8	0,44
Model orientat a objectes	1,09	1,06	20 de 30	0,33

Podem convenir doncs que la correspondència més propera a 1 entre sèries i sistemes d'informació es dona en el model orientat a objectes, probablement pel pes que s'ha donat a la gestió d'expedients electrònics en el cas d'aplicació del Consorci. Els conjunts de dades simples (model *data lake*) també reflecteixen de forma clara la seva pertinença a una única sèrie documental, tot i que sense caràcter exhaustiu, atès que la majoria de sèries utilitzen més d'un conjunt de dades.

En el model relacional la proporció s'inverteix, com a resultat probablement del cost que representa desenvolupar una aplicació informàtica amb base relacional, cosa la qual porta a agrupar tràmits (i per tant sèries documentals) dins el funcionament d'una única aplicació.

L'índex de coincidència en dades dona tanmateix valors en general baixos, inferiors al 0,5, en la correspondència 1 a 1 entre sèrie i sistema, cosa la qual indica que en pocs casos el total de dades de la sèrie es pot gestionar únicament amb el sistema d'informació (és a dir, la sèrie està poc dadificada), i encara es depèn en un percentatge important d'evidències-objecte.

7.5 Objectiu 4: les arquitectures distribuïdes com a entorns de gestió (transaccionals)

La pregunta que es planteja és si els sistemes distribuïts tipus “*data fabric*”, que combinen l’ús de conjunts de dades simples (“*data lake*”), habitualment amb origen ofimàtic, i de petites solucions en entorns de col·laboració (desenvolupaments *low code* en sistemes orientats a objectes) són els entorns que tenen actualment més pes en la gestió diària dels processos, i per tant es poden considerar el principal entorn transaccional (per al concepte de *data fabric* vegeu apartat 2.2.3.5 del capítol 2).

En aquest sentit, considerarem “transaccionals” aquells sistemes que tenen com a finalitat principal la gestió de processos, de forma parcial o completa, i seguint un determinat curs en la seva tramitació. En oposició a aquesta funció, parlarem de “sistemes registrals”, o del caràcter “registrar” d’un sistema d’informació, quan la principal funcionalitat del sistema sigui deixar constància de l’estat en un moment donat d’una cosa o d’una acció.²⁸¹ En aquest sentit, els sistemes registrals mantindran les seves dades únicament amb processos d’alta, baixa i modificació, però sense cap flux de tramitació predeterminat.

Explorem doncs els indicadors que considerem que ens poden identificar un sistema d’informació com a transaccional.

Per això utilitzarem en primer lloc l’indicador **SIS-05**, que correspon al **nombre de processos que l’utilitzen**. La hipòtesi és que els sistemes transaccionals estaran relacionats majoritàriament amb un sol procés, mentre que els sistemes que gestionen dades mestres, pel seu caràcter de subministradors d’informació, estaran relacionats amb un nombre significatiu de processos.

Una primera observació dels deu sistemes d’informació amb més processos vinculats ens dona la següent relació:

Taula 7-38: Sistemes d’informació amb més processos vinculats

SISTEMA D'INFORMACIÓ	TECNOLOGIA	PROCESSOS (SIS-05)
SI04.008 GEDAC		7
SI01.029 Registre del Pressupost		4
SI04.021 Registre d'alumnes de Catalunya (RALC)		4
SI02.012 Catàleg de processos del CEB		4

²⁸¹ Utilitzarem el terme registrar, tant referit a sistemes en el seu conjunt com a funcionalitats dels sistemes, de forma equivalent al de gestió de dades mestres (per a més detall vegeu l’apartat 2.2.2.1.2 del capítol 2).

SISTEMA D'INFORMACIÓ	TECNOLOGIA	PROCESSOS (SIS-05)
SI02.001 Base de Dades de Centres (BDC)		3
SI05.003 GIP-SIP		3
SI04.023 Gestió dels centres docents (CDE)		3
SI04.024 Gestió del pagament delegat (Nòmina, Seguretat social i IRPF) (GPD)		3
SI01.008 Registre del dipòsit de materials		2
SI01.009 Registre de seguiment dels informes d'inspecció		2

Observem una característica tecnològica comuna, i és que dels deu sistemes amb més processos vinculats, set són aplicacions informàtiques convencionals, basades en bases de dades relacionals, i tres ofimàtica. No apareix entre els deu primers cap desenvolupament *low code* basat en gestors documentals. Anàlitzem doncs, prèvia transformació a valors numèrics dels models d'emmagatzemament i de desenvolupament, la possible correlació, per al conjunt del mapa de sistemes i processos, entre la tecnologia utilitzada i el nombre de processos vinculats. El coeficient de correlació de Pearson dona el següent resultat:

Taula 7-39: Coeficient de correlació del nombre de processos i els models tecnològics

	MODEL D'EMMAGATZEMAMENT	MODEL DE DESENVOLUPAMENT
Nombre de processos per sistema (SIS-05)	0,29	0,20

La correlació observada és massa baixa per a ser considerada significativa. Ens remetem doncs a una mètrica més simple, les mitjanes per cada model, per analitzar la relació entre el model de desenvolupament i emmagatzemament i el nombre de processos:

Taula 7-40: Processos per sistema d'informació per model d'emmagatzemament

MODEL D'EMMAGATZEMAMENT	PROCESSOS PER SISTEMA
Base de dades relacional	1,57
Conjunts de dades simples	1,08
Base de dades orientada a objectes (GD)	1,04

Taula 7-41: Processos per sistema d'informació per model de desenvolupament

MODEL DE DESENVOLUPAMENT	PROCESSOS PER SISTEMA
Desenvolupament convencional	1,54
Sense desenvolupament	1,11
Desenvolupament <i>low code</i>	1,04

Si ho complementem amb l'observació del caràcter funcional dels deu sistemes amb més vincles amb processos observem que un 80% correspon a sistemes que fan funcions registrals, i que per tant mantenen, o són candidats a mantenir, dades mestres rellevants per al Consorci.

Taula 7-42: Caràcter estimat dels sistemes d'informació vinculats a més processos

SISTEMA D'INFORMACIÓ	TIPUS DE DADES	CARÀCTER ESTIMAT
SI04.008 GEDAC		Transaccional
SI01.029 Registre del Pressupost		Transaccional
SI04.021 Registre d'alumnes de Catalunya (RALC)		Funcions de registre d'alumnat
SI02.012 Catàleg de processos del CEB		Funcions de registre d'actius d'informació
SI02.001 Base de Dades de Centres (BDC)		Funcions de registre de centres educatius
SI05.003 GIP-SIP		Funcions de registre de personal
SI04.023 Gestió dels centres docents (CDE)		Funcions de registre de centres educatius
SI04.024 Gestió del pagament delegat (Nòmina, Seguretat social i IRPF) (GPD)		Funcions de registre de personal concertat
SI01.008 Registre del dipòsit de materials		Funcions de registre de mobiliari reciclat
SI01.009 Registre de seguiment dels informes d'inspecció		Funcions de registre d'informes

Podem convenir en conseqüència que el nombre de processos als quals dona suport (indicador SIS-05) és un indicador possiblement vàlid del caràcter del sistema com a sistema de caràcter registral, o gestor de dades mestres, però no ens confirma que en valors baixos ens determini la transaccionalitat del sistema.

La segona opció a valorar és el **nombre d'evidències que genera un sistema** (indicador **SIS-07**). La hipòtesi és que un sistema transaccional generarà moltes evidències en un sol procés i sèrie, mentre que un sistema registral generarà poques evidències en molts processos i sèries diferents.

Ho avaluarem utilitzant el grup focal. Triem doncs els sistemes del grup focal que tenen relació amb més d'un procés i sèrie documental, i observem la quantitat d'evidències amb què es vinculen. Són els sistemes que en teoria haurien de contenir dades mestres.

Taula 7-43: Nombre d'evidències per sistema d'informació dels sistemes vinculats a més sèries

SISTEMA D'INFORMACIÓ	SÈRIE DOCUMENTAL	EVIDÈNCIES	CARÀCTER ESTIMAT
SI02.012 Catàleg de processos del CEB	INF001-001 Explotació de dades	1	Funcions de registre d'actius d'informació
	INF001-002 Manteniment del banc de dades del CEB	2	
	ORG003-001 Anàlisi i millora de processos	1	
SI05.034 GEAC (Aplicació de Gestió Expedients de Arxiu Central)	DOC003-001 Transferències a l'Arxiu Central	1	Transaccional amb funcions de registre d'expedients
	DOC004-001 Préstecs i consultes de documentació	6	
SI07.014 Àrea de client. Gestió de préstec i consultes de dipòsit documental	DOC003-001 Transferències a l'Arxiu Central	1	Transaccional amb funcions de registre d'expedients
	DOC004-001 Préstecs i consultes de documentació	2	
SI08.012 SDOC001-002 Assessorament als centres i serveis educatius en matèria de gestió documental	DOC001-001 Desenvolupament del sistema de gestió documental	1	Transaccional
	DOC001-002 Assessorament als centres en matèria de gest. doc.	4	

No observem cap diferència significativa entre els sistemes transaccionals i registrals. No podem inferir, per tant, cap relació entre el nombre d'evidències i el caràcter del sistema d'informació. Això ens porta a descartar l'indicador SIS-07 com a possible indicador del caràcter transaccional d'un sistema.

La tercera opció és pel marc de **relacions del sistema amb altres sistemes d'informació**. Per avaluar aquesta variable disposem de tres indicadors que descriuen el rol del sistema en el conjunt del mapa de sistemes, els quals ens permeten construir una hipòtesi combinada. La **hipòtesi** és la següent:

Taula 7-44: Hipòtesi sobre les relacions entre sistemes d'informació

MESURA O INDICADOR	Un sistema d'informació transaccional es caracteritza per...	Un sistema d'informació registral es caracteritza per...
SIS-13 Índex de criticitat del sistema	Un nombre limitat de sistemes amb els quals es relaciona.	Un nombre elevat de sistemes amb els quals es relaciona.
SIS-14 Índex d'influència del sistema	Un índex d'influència que tendeix a 0 (equilibri entre les entrades i les sortides)	Un índex d'influència superior a 0 (majoritàriament proporciona dades a altres sistemes)
SIS-15 Índex de connectivitat ponderada del sistema	Un índex de connectivitat baix (poca automatització en les integracions amb altres sistemes)	Un índex de connectivitat alt, que denota connexions directes o automatitzades amb altres sistemes

Hem realitzat una anàlisi de conglomerats jeràrquica utilitzant aquests tres indicadors sobre tots els sistemes d'informació vigents que tenen algun vincle amb altres sistemes. El resultat han estat tres *clústers* de dimensions diferents. El primer *clúster* agrupa set sistemes amb les següents característiques:

Taula 7-45: Models tecnològics dels sistemes d'informació del primer clúster

SISTEMA D'INFORMACIÓ	TECNOLOGIA	MODEL DE DESENVOLUP.	MODEL D'EMMAGAT.	CARÀCTER ESTIMAT
SI04.002 Esfer@				Transaccional amb funcions de registre d'alumnat
SI04.008 GEDAC				Transaccional
SI04.009 Gestió d'usuaris d'altres centres (GUAC)				Funcions de registre d'usuaris
SI04.013 POE Planificació Oferta Educativa				Funcions de registre de l'oferta educativa
SI04.021 Registre d'alumnes de Catalunya (RALC)				Funcions de registre d'alumnat
SI04.037 Taules Mestres del departament (TME)				Taules mestres del Dept. d'Educ.
SI04.043 Sistema analític d'indicadors corporatius (SIDIE)				Sistema d'explotació

En 5 dels 7 casos es tracta de sistemes d'informació de grans dimensions que contenen dades mestres. El segon *clúster* agrupa dos sistemes:

Taula 7-46: Models tecnològics dels sistemes d'informació del segon clúster

SISTEMA D'INFORMACIÓ	TECNOLOGIA	MODEL DE DESENVOLUP.	MODEL D'EMMAGAT.	CARÀCTER ESTIMAT
SI04.012 OFE-EDU				Funcions de registre de l'oferta educativa
SI10.001 Alumnat NEE-A				Sistema d'exploració

No hem observat diferències significatives entre el primer i el segon *clústers*.

La resta de sistemes d'informació (172) corresponen al tercer *clúster*, i es distribueixen de la forma següent:

Taula 7-47: Models tecnològics dels sistemes d'informació del tercer clúster

	DESENVOL. CONVENCIONAL	DESENVOL. LOW CODE	SENSE DESENVOL.
Base de dades orientada a objectes (GD)	0	83	0
Base de dades relacional	22	0	0
Conjunts de dades simples	0	24	38
Data warehouse (BI)	5	0	0

Dins aquest tercer *clúster* trobem inclosos també nombrosos sistemes que gestionen dades mestres:

Taula 7-48: Sistemes d'informació del tercer clúster que gestionen dades mestres

SISTEMA D'INFORMACIÓ	TECNOLOGIA	MODEL DE DESENVOLUP.	MODEL D'EMMAGAT.
SI01.236 Registre de programes d'innovació / educatius del CEB			
SI02.015 Registre d'Alumnat Vulnerable (RAV)			
SI08.048 EDU-1659-L1 Registre d'inspectors			
SI08.051 GPC005-002-L1 Registre de programes CUP/CUC			

No podem considerar aquests tres indicadors combinats com a determinants del caràcter transaccional o registral d'un sistema d'informació. Això no exclou que valors elevats, associats sobretot a tecnologies més robustes, puguin ser utilitzats com un indicador complementari de sistema gestor de dades mestres.

Utilitzem a continuació la quarta modalitat de l'índex d'influència, que correspon a l'indicador **SIS-16 Índex de connectivitat potencial del sistema**. Aquest indicador valora quin nivell de coincidència tenen les dades d'un sistema amb les dades d'altres sistemes. Però també hi inclouem una variant, el nombre de sistemes d'informació amb els quals hi ha coincidència. D'acord amb això, podem inferir com a hipòtesi de partida que un sistema registral tindrà coincidència amb un nombre elevat de sistemes, i per tant un potencial elevat de connectivitat i de reutilització de la seva informació en altres contextos, sense que el nombre de dades coincidents hagi de ser necessàriament alt.

Si observem els cinc sistemes amb un major nombre de coincidències, obtenim:

Taula 7-49: Sistemes d'informació amb un major nombre de coincidències de dades amb altres sistemes

SISTEMA D'INFORMACIÓ	MODEL DE DESENVOLUP.	MODEL D'EMMAGAT.	SISTEMES AMB COINCIDÈNCIA	ÍNDEX DE COINCIDÈNCIA (SIS-16) ²⁸²
SI04.002 Esfer@			211	14,14
SI04.164 Col·laboració professional no docent en horari lectiu			203	13,93
SI04.009 Gestió d'usuaris d'altres centres (GUAC)			198	15,68
SI04.022 Gestió dels centres privats amb ensenyaments concertats (GConcert)			197	16,75
SI05.039 Reclamacions patrimonials			197	17,01

²⁸² El valor indicat és la mitjana dels índex SIS-16 de cada sistema amb la resta de sistemes amb els quals té alguna dada coincident.

No observem tampoc una regularitat clara, atès que hi trobem sistemes transaccionals purs (SI05.039 i SI04.164) i sistemes transaccionals que fan funcions de registre (SI04.002, SI04.009 i SI04.022). Si observem a continuació els sistemes amb un índex de connectivitat potencial més alt, els resultats són aquests:

Taula 7-50: Sistemes d'informació amb un índex de coincidència major

SISTEMA D'INFORMACIÓ	MODEL DE DESENVOLUP.	MODEL D'EMMAGAT.	SISTEMES AMB COINCIDÈNCIA	ÍNDEX DE COINCIDÈNCIA (SIS-16)
SI04.047 Comunicació absentisme dels centres educatius			151	35,26
SI01.124 Usuaris servei de menjador			151	35,26
SI01.035 Repositori de contractes			20	34,71
SI01.193 Eina seguiment estat de la licitació			20	34,71
SI01.198 Inventari de mobiliari nou (catàleg)			30	32,89

Novament no observem cap regularitat que ens suggereixi que aquest indicador pot ser discriminant en la determinació del caràcter del sistema d'informació: dos sistemes són transaccionals (SI04.047 i SI01.198) i tres podrien considerar-se candidats a gestors de dades mestres (SI01.124, SI01.035 i SI01.198). Tot i així, i considerant que el sistema SI04.047 també té un cert component registral, valorem en el futur continuar amb l'anàlisi d'aquest indicador.

I la quarta opció és per l'índex de transaccionalitat, en les seves tres modalitats: a partir de les dades elementals, a partir del tipus de dades, i a partir dels objectes (indicadors **SIS-17**, **SIS-18** i **SIS-19**). La hipòtesi és que un sistema transaccional contindrà més dades descriptives d'objectes del tipus "acció", i més dades de tipus transaccional, i que un sistema de dades mestres contindrà un nombre elevat de dades de pocs objectes del tipus "cosa", i hi predominaran les dades descriptives.

Els cinc sistemes amb un índex de transaccionalitat basat en les dades elementals (SIS-17) més alt són els següents:

Taula 7-51: Sistemes d'informació amb un índex de transaccionalitat basat en dades més alt

SISTEMA D'INFORMACIÓ	INDEX SIS-17	CARÀCTER ESTIMAT
SI08.024 SLOP001-001 Desenvolupament del sistema de protecció de dades de caràcter personal	1	Transaccional
SI08.025 SLOP007-001 Assessorament en matèria de protecció de dades personals	1	Transaccional
SI08.028 SLOP005-001 Auditories APDCAT i controls de seguretat sobre protecció de dades personals	1	Transaccional
SI08.019 SGRH018-002 Gestió dels recursos humans (USIDP)	1	Transaccional amb funcions de registre de personal
SI08.014 DOC003-001-L1 Transferències a Arxiu Central	1	Transaccional

Observem doncs que són en tots els casos clarament transaccionals, sense excloure que també gestionin en algun cas dades mestres. La mateixa observació amb els cinc sistemes amb un índex de transaccionalitat basat en el tipus de dades (SIS-18) més alt són els següents:

Taula 7-52: Sistemes d'informació amb un índex de transaccionalitat basat en el tipus de dades més alt

SISTEMA D'INFORMACIÓ	INDEX SIS-18	CARÀCTER ESTIMAT
SI08.024 SLOP001-001 Desenvolupament del sistema de protecció de dades de caràcter personal	0,75	Transaccional
SI08.025 SLOP007-001 Assessorament en matèria de protecció de dades personals	0,75	Transaccional
SI08.028 SLOP005-001 Auditories APDCAT i controls de seguretat sobre protecció de dades personals	0,75	Transaccional
SI08.026 SLOP002-001 Atenció al dret d'accés a les dades personals, de rectificació de dades personals, [...]	0,71	Transaccional
SI08.019 SGRH018-002 Gestió dels recursos humans (USIDP)	0,60	Transaccional amb funcions de registre de personal

Novament observem que es confirma el caràcter transaccional dels sistemes. I la mateixa observació amb els cinc sistemes amb un índex de transaccionalitat basat en objectes (SIS-19) més alt són els següents:

Taula 7-53: Sistemes d'informació amb un índex de transaccionalitat basat en objectes més alt

SISTEMA D'INFORMACIÓ	INDEX SIS-19	CARÀCTER ESTIMAT
SI08.024 SLOP001-001 Desenvolupament del sistema de protecció de dades de caràcter personal	1,00	Transaccional
SI08.025 SLOP007-001 Assessorament en matèria de protecció de dades personals	1,00	Transaccional
SI08.028 SLOP005-001 Auditories APDCAT i controls de seguretat sobre protecció de dades personals	1,00	Transaccional
SI08.019 SGRH018-002 Gestió dels recursos humans (USIDP)	1,00	Transaccional amb funcions de registre de personal
SI08.014 DOC003-001-L1 Transferències a Arxiu Central	1,00	Transaccional

Per tant, el caràcter transaccional dels sistemes pot venir determinat pels indicadors de transaccionalitat. Però cal confirmar-ho en la franja oposada, els valors més baixos dels índex de transaccionalitat. Analitzant deu sistemes clarament identificables com a candidats a sistemes registrals, observem doncs que tots ells se situen en la franja més baixa dels valors d'aquest indicador:

Taula 7-54: Comparativa dels índex de transaccionalitat dels sistemes de caràcter registral

SISTEMA D'INFORMACIÓ	INDEX SIS-17	INDEX SIS-18	INDEX SIS-19
SI01.014 Base de dades del personal subaltern municipal	0,00	0,00	0,00
SI01.044 Registre centres	0,00	0,00	0,00
SI01.097 Registre alumnat convocatòria	0,00	0,00	0,00
SI01.106 Autoritzacions, modificacions i cessaments d'estudis	0,00	0,00	0,00
SI01.124 Usuaris servei de menjador	0,00	0,00	0,00
SI02.001 Base de Dades de Centres (BDC)	0,00	0,00	0,00
SI04.009 Gestió d'usuaris d'altres centres (GUAC)	0,03	0,00	0,14
SI01.061 Base de dades UEC	0,07	0,14	0,25
SI04.021 Registre d'alumnes de Catalunya (RALC)	0,11	0,08	0,33
SI01.036 BBDD mare (edificis)	0,12	0,06	0,33

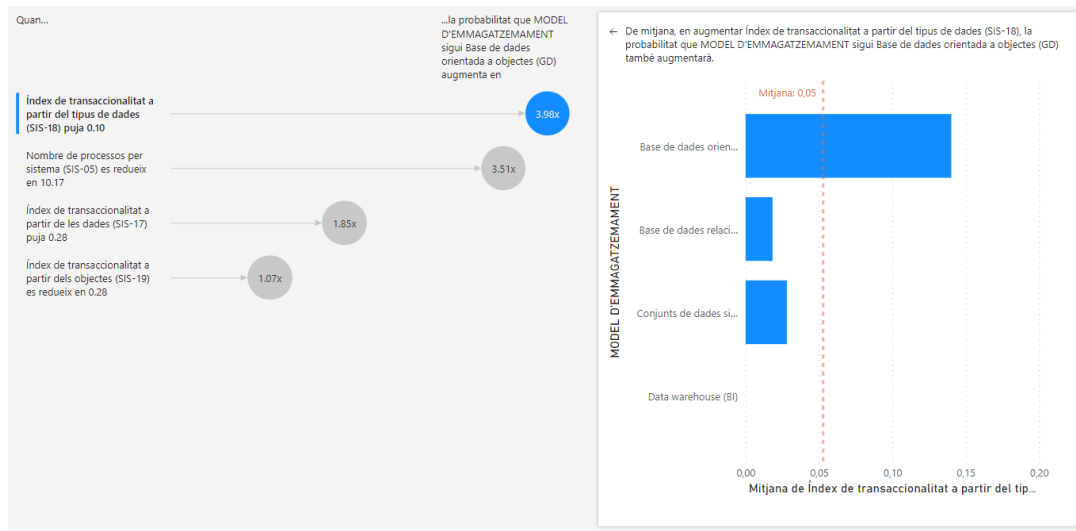
Podem convenir doncs que les tres modalitats de l'índex de transaccionalitat són indicadors rellevants per a la classificació dels sistemes com a transaccionals.

Finalment hem utilitzat els dos indicadors que en l'estudi exploratori han donat mostres de poder determinar la transaccionalitat d'un sistema d'informació, el nombre de processos vinculats (indicador SIS-05) i els tres índex de transaccionalitat (indicadors

SIS-17, SIS-18 i SIS-19), per a determinar la tecnologia que s'associa majoritàriament als sistemes transaccionals, amb l'objectiu de confirmar o refutar la hipòtesi de l'objectiu 4. Per a fer-ho hem utilitzat la funcionalitat de càlcul dels influenciadors clau de *Microsoft Power BI*. Els resultats per indicador han estat els que es detallen a continuació.

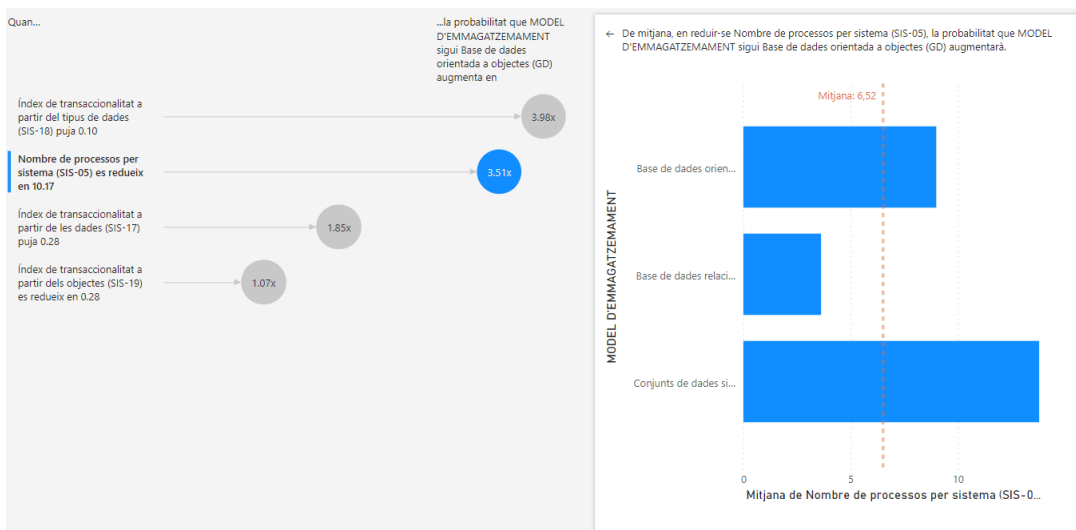
Aplicant l'índex de transaccionalitat a partir del tipus de dades (SIS-18):

Figura 7-24: Nivell d'influència de l'índex de transaccionalitat basat en el tipus de dades sobre el model tecnològic



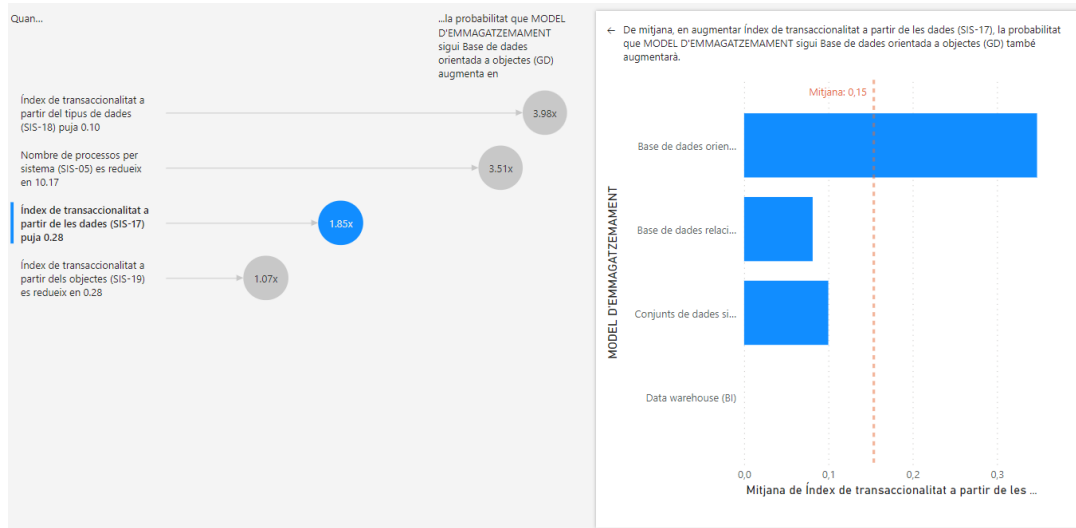
Aplicant el nombre de processos per sistema (SIS-05):

Figura 7-25: Nivell d'influència del nombre de processos sobre el model tecnològic



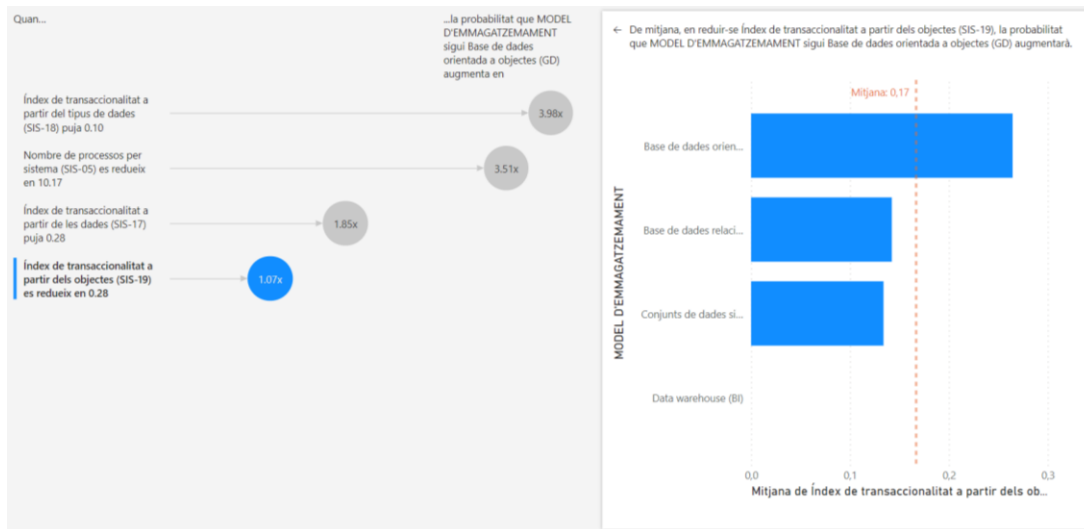
Aplicant l'índex de transaccionalitat a partir de les dades (SIS-17):

Figura 7-26: Nivell d'influència de l'índex de transaccionalitat basat en dades sobre el model tecnològic



Aplicant l'índex de transaccionalitat a partir dels objectes (SIS-19):

Figura 7-27: Nivell d'influència de l'índex de transaccionalitat basat en objectes sobre el model tecnològic



7.5.1 Conclusió

A partir de les anàlisi realitzades, s'ha observat que el nombre de processos i els índex de transaccionalitat són indicadors rellevants per a la classificació dels sistemes com a transaccionals. El nombre d'evidències, les relacions del sistema amb altres sistemes d'informació i l'índex de connectivitat potencial del sistema no es consideren tanmateix indicadors significatius per assignar el caràcter de transaccional a un sistema d'informació.

Amb aquests dos indicadors, s'ha pogut determinar que, en la línia d'evolució tecnològica que ha seguit el Consorci d'Educació de Barcelona durant els darrers quatre anys, el model d'emmagatzematge distribuït tipus *data lake* o *data fabric*, combinant desenvolupaments *low code* en entorns de col·laboració orientats a objectes amb conjunts de dades simples, s'ha convertit en l'entorn que concentra el gruix de la gestió del Consorci, atès que respon al nivell general de transaccionalitat més alt.

7.6 Objectiu 5: equivalència entre tipus de sèries documentals i tipus de sistemes d'informació

Un cop classificats els sistemes d'informació i categoritzades les sèries documentals, volem esbrinar si existeix alguna relació significativa entre el tipus de sèrie documental i el tipus de sistema d'informació. La hipòtesi és que l'equivalència entre sèrie documental i sistema d'informació ve determinada per la tipologia de cadascuna de les dues entitats. Així doncs, les sèries documentals de tramitació són equivalents a sistemes d'informació transaccionals, mentre que les sèries documentals recapitulatives s'associen a sistemes registrals.

Per avaluar aquest objectiu treballarem amb dos tipus de categories per classificar les sèries documentals del Consorci. La primera és la que distingeix entre sèries documentals "naturals", és a dir emanades directament de l'activitat (i en conseqüència purament funcionals), que en aquest context anomenarem transaccionals, i les sèries documentals "artificials", que es poden construir per acumulació de fraccions de sèries transaccionals (les sèries recapitulatives o registrals) o per agrupació d'evidències en funció de la temàtica o qualsevol altre aspecte formal (les sèries factícies).²⁸³

La segona categoria tindrà en compte el tipus de tramitació o procediment de l'activitat documentada, i en funció del coneixement de les dates extremes i de la previsibilitat del contingut, parlarem de sèries reglades, tècniques o temàtiques.

Prèviament a posar en relació les sèries documentals amb els sistemes d'informació, analitzarem si disposem d'indicadors que en puguin determinar alguna de les dues categories esmentades. Partint de considerar òptim que una sèrie documental sigui transaccional i reglada, enumerem doncs a continuació les variables que analitzarem si poden determinar el caràcter transaccional d'una sèrie documental.

El primer indicador serà la **relació de la sèrie amb altres sèries**, que trobem representat en les tres modalitats de l'índex d'influència: simple (SER-10), ponderat (SER-11) i net (SER-12). Ens basarem per a l'ordenació en l'índex d'influència net, que té en compte també la influència del procés. Analitzarem el caràcter de les deu sèries primàries vigents amb un índex d'influència net major, i el compararem amb les deu que tenen un índex d'influència menor.²⁸⁴ A partir del resultat, valorarem el caràcter discriminant d'aquest indicador.

²⁸³ Sobre les tipologies de sèries vegeu l'apartat 2.3.1.4.1 de l'estat de la qüestió (capítol 2).

²⁸⁴ No prenem en consideració les sèries secundàries, atès que en el procés de dadificació aquestes desapareixen, i per tant no són rellevants als efectes d'equivalència amb sistemes d'informació.

Taula 7-55: Sèries documentals amb un índex d'influència més alt i més baix

SÈRIE DOCUMENTAL	SER-10	SER-11	SER-12	CARÀCTER	TIPUS
GRH016-002 Jubilacions	1	2	28	Transaccional	Reglada
GRH016-004 Expedients contradictoris	1	2	28	Transaccional	Reglada
GRH016-001 Relació contractual i ocupació de llocs	3	-2	26	Transaccional	Reglada
ORG003-001 Anàlisi i millora de processos	0	0	26	Transaccional	No reglada tècnica
GRH016-003 Avaluació de personal	1	-2	24	Transaccional	Reglada
ECO001-001 Gestió de la despesa	7	14	20	Factícia	No reglada temàtica
INF001-002 Manteniment del banc de dades del CEB	6	11	20	Transaccional	Reglada
NEE008-004 Protocol per avaluar les necessitats de suport complementari	1	2	19	Transaccional	Reglada
ESC005-002 Preinscripció, matrícula i nova admissió	6	8	18	Transaccional	Reglada
TIC007-001 Desenvolupament de solucions TIC	2	3	17	Transaccional	Reglada
Mitjana			4		
ATU004-001 Llistats de distribució del registre d'entrada	1	-2	-2	Factícia	No reglada temàtica
ATU004-002 Albarans d'entrega de correspondència a l'empresa distrib.	1	-2	-2	Factícia	No reglada temàtica
ECO001-003 Gestió de la despesa	1	-2	-2	Factícia	No reglada temàtica
ECO001-007 Gestió de la despesa de postob.	1	-2	-2	Factícia	No reglada temàtica
ECO001-010 Habilitació	1	-2	-2	Transaccional	No reglada tècnica
ESC011-001 Reclamacions de preinscripció dels estudis postob.	1	-2	-2	Transaccional	Reglada
ESC011-002 Validació de NEE en el procés d'admissió als ens. postob.	1	-2	-2	Transaccional	Reglada
NEE010-001 Sol·licituds de preins., reclam. i nova admissió en CEE ²⁸⁵	0	0	-3	Transaccional	Reglada
ORG004-002 Normalització de formularis i models de document	1	-2	-3	Transaccional	No reglada tècnica
LOP005-001 Auditories APDCAT i controls de seguretat sobre prot. [...]	0	0	-4	Transaccional	No reglada tècnica

En l'anàlisi de les deu sèries extremes no s'identifica cap regularitat determinant. El nombre de sèries documentals transaccionals és superior en les que obtenen un índex

²⁸⁵ Centres d'educació especial.

d'influència net més alt, però no exclou aquesta categoria en els valors més baixos de l'índex. Sí que s'aprecia diferència en el caràcter reglat o no reglat de la sèrie documental, essent majoritàries les sèries reglades quan d'índex d'influència és superior. No s'ha identificat cap sèrie documental que es pugui considerar recapitulativa, però sí sèries factícies o de conveniència, majoritàriament amb un índex d'influència baix. No podem per tant determinar el valor de l'índex d'influència com a predictor del caràcter transaccional de la sèrie documental.

El segon indicador que plantegem és el **nombre d'evidències de la sèrie documental** (indicador **SER-02**). Una sèrie documental transaccional, com a resultat que és de l'execució d'una sèrie d'accions o tasques consecutives, generarà un nombre mínim d'evidències, que en el cas del Consorci oscil·len entre 5 i 6 per sèrie. Però desconeixem el nombre mitjà d'evidències que pot generar una sèrie recapitulativa o factícia, i si la diferència és prou significativa com per considerar vàlid aquest indicador. Observem les dades de les sèries del grup focal amb un nombre menor d'evidències.

Taula 7-56: Sèries documentals del grup focal amb un nombre menor d'evidències

SÈRIE DOCUMENTAL	EVIDÈNCIES	CARÀCTER	TIPUS
INF007-001 Coordinació de l'ús social de centres	2	Transaccional	No reglada tècnica
ORG004-002 Normalització de formularis i models de document	2	Transaccional	No reglada tècnica
TRA001-001 Publicació d'informació a l'espai de transparència	2	Transaccional	No reglada tècnica
TRA003-001 Assessorament en matèria de transparència als centres educatius	2	Transaccional	No reglada tècnica
INF002-001 Eleccions al Consell Escolar de Centre	1	Transaccional	Reglada
INF009-002 Manteniment de les dades de la Base de dades de Centres (BDC)	1	Recapitulativa	---
NEE020-001 Intervencions en matèria de seguretat ciutadana i prevenció [...]	1	Transaccional	No reglada tècnica

Llevat de la sèrie INF009-002, que es podria considerar recapitulativa per la naturalesa del sistema d'informació que manté, en la resta de casos són sèries associades a activitats, i que es poden considerar transaccionals. El motiu del baix nombre d'evidències es deu a que es tracta d'activitats de pregunta-resposta (TRA003-001) o de mera comunicació unidireccional (INF007-001, INF002-001 o NEE020-001), i per aquest motiu de caràcter no reglat tècnic. No es considera per tant que l'indicador SER-

02 sigui determinant per a la categorització de la sèrie documental, atès que sèries clarament transaccionals poden donar índex inferiors a la mitjana observada al Consorci.

Analitzem un tercer indicador, que és l'índex de dadificació simple a nivell d'evidència (indicador **SER-13**). Inferim d'inici que una sèrie documental recapitulativa dadificada es basarà fortament en dades, i per tant la major part d'evidències que la composin seran de tipus "dades". Analitzem doncs, de les sèries documentals del grup focal, les quinze sèries amb un major índex de dadificació basat en evidències.

Taula 7-57: Sèries documentals del grup focal amb un major índex de dadificació basat en evidències

SÈRIE DOCUMENTAL	ÍNDEX DE DADIFICACIÓ	CARÀCTER	TIPUS
GRH023-001 Gestió de borses de personal de suport als centres educatius (COVID-19)	1,00	Transaccional	Reglada
INF002-001 Eleccions al Consell Escolar de Centre	1,00	Transaccional	Reglada
INF007-001 Coordinació de l'ús social de centres	1,00	Transaccional	No reglada tècnica
NEE020-001 Intervencions en matèria de seguretat ciutadana i prevenció de radicalismes violents	1,00	Transaccional	No reglada tècnica
TRA001-001 Publicació d'informació a l'espai de transparència	1,00	Transaccional	No reglada tècnica
NEE002-001 Actuacions en casos de sospita de maltractament i/o assetjament en un centre educatiu	0,83	Transaccional	No reglada tècnica
DOC001-002 Assessorament als centres i serveis educatius en matèria de gestió documental	0,80	Transaccional	No reglada tècnica
INF001-002 Manteniment del banc de dades del CEB	0,80	Transaccional	Reglada
GRH012-004 Comunicació d'exercici del dret de vaga	0,67	Transaccional	Reglada
INF001-001 Explotació de dades	0,63	Transaccional	Reglada
INF009-001 Manteniment de les dades de la Base de Dades de Centres (BDC)	0,60	Recapitulativa	No reglada tècnica
LOP007-001 Assessorament en matèria de protecció de dades personals	0,60	Transaccional	No reglada tècnica

SÈRIE DOCUMENTAL	ÍNDEX DE DADIFICACIÓ	CARÀCTER	TIPUS
NEE021-001 Seguiment i valoració d'alumnes en situació de vulnerabilitat i/o malestar social o psicosocial	0,50	Transaccional	Reglada
ORG004-001 Normalització de formularis i models de documents	0,45	Transaccional	No reglada tècnica
ORG003-001 Anàlisi i millora de processos	0,44	Transaccional	No reglada tècnica

Novament observem que la única sèrie candidata a recapitulativa identificada entre les quinze sèries amb un índex de dadificació més alt no presenta cap singularitat en la seva mesura. Pel que fa al tipus de sèries, tampoc s'aprecia cap regularitat. Les sèries amb menys documents, quan aquests documents són formularis electrònics de comunicació, sí que presenten valors més alts en l'índex de dadificació. Descartem doncs aquests indicador com a discriminador del caràcter de la sèrie.

Proposem un quart indicador que és el **tipus d'objectes descrits** (basat en l'indicador **SER-08**). La hipòtesi és que una sèrie transaccional descriurà preferentment un objecte del tipus "acció" i diversos objectes del tipus "cosa", mentre que una sèrie recapitulativa només descriurà objectes del tipus "cosa". En aquest cas, les dues úniques sèries documentals del grup focal clarament candidates a recapitulatives compleixen el patró de no descriure cap objecte de tipus "acció", sinó només de tipus cosa.

Taula 7-58: Sèries documentals del grup focal candidates a sèries recapitulatives

SÈRIE DOCUMENTAL	ACCIÓ	COSA
INF009-001 Manteniment de les dades de la Base de Dades de Centres (BDC)	0	7
INF009-002 Manteniment de les dades de la Base de dades de Centres (BDC)	0	3

La resta de sèries totes descriuen entre un i quatre objectes de tipus "acció". Podem considerar per tant aquesta una variable vàlida per a discriminar el caràcter d'una sèrie documental. Si introduïm com a variable el tipus d'expedient generat, podem obtenir una representació més completa d'aquest indicador:

Taula 7-59: Mitjana d'objectes per tipus d'objecte i tipus de sèrie documental

CARÀCTER I TIPUS	MITJANA OBJECTES ACCIÓ	MITJANA OBJECTES COSA	PROPORCIÓ D'OBJECTES ACCIÓ
Recapitulativa	0,00	5,00	0%
Transaccional - No reglada tècnica	1,56	3,13	35%
Transaccional - Reglada	2,14	4,71	30%

Observem que, en general, les sèries amb tramitació reglada impliquen un nombre major d'objectes que les sèries no reglades, tot i que la proporció dels objectes tipus "acció" sobre el total és similar en ambdós casos, sobre el 30-35% del total d'objectes. Valorem positivament doncs la utilitat d'aquest indicador com a possible discriminador del caràcter de la sèrie documental, i també del tipus.

Aquesta constatació ens porta a incorporar com a variable l'índex de transaccionalitat de la sèrie, en les seves tres modalitats: Índex de transaccionalitat a partir de les dades elementals (SER-18), Índex de transaccionalitat a partir del tipus de dades (SER-19) i Índex de transaccionalitat a partir dels objectes (SER-20), que és una evolució de l'anterior indicador SER-08. S'ha incorporat també una anàlisi *clúster* per valorar la correspondència entre els valors dels tres índex de transaccionalitat i els tipus de sèrie documental. Els valors d'aquest indicador per a les 32 sèries documentals del grup focal són els següents:

Taula 7-60: Valors dels índex de transaccionalitat per a les sèries documentals del grup focal

SÈRIE DOCUMENTAL	ÍNDEX SER-18	ÍNDEX SER-19	ÍNDEX SER-20	CLUSTER
DOC001-001 Desenvolupament del sistema de gestió documental	0,63	0,50	0,33	Clúster1
DOC001-002 Assessorament als centres i serveis educatius en matèria de gestió documental	0,54	0,38	0,25	Clúster1
DOC002-001 Tractament documental	0,60	0,33	0,40	Clúster1
DOC003-001 Transferències a l'Arxiu Central	0,45	0,18	0,30	Clúster1
DOC004-001 Préstecs i consultes de documentació	0,45	0,21	0,33	Clúster1
ESC001-001 Seguiment de falsos empadronaments	0,31	0,22	0,27	Clúster1
GRH012-004 Comunicació d'exercici del dret de vaga	0,47	0,12	0,43	Clúster1
GRH023-001 Gestió de borses de personal de suport als centres educatius (COVID-19)	0,32	0,20	0,25	Clúster1
INF001-001 Explotació de dades	0,50	0,34	0,17	Clúster1
INF001-002 Manteniment del banc de dades del CEB	0,79	0,07	0,40	Clúster1
INF001-003 Elaboració d'estudis amb col·laboracions externes	0,53	0,18	0,20	Clúster1
INF002-001 Eleccions al Consell Escolar de Centre	0,00	0,06	0,00	Clúster2
INF007-001 Coordinació de l'ús social de centres	0,64	0,18	0,40	Clúster1
INF009-001 Manteniment de les dades de la Base de Dades de Centres (BDC)	0,00	0,07	0,00	Clúster2
INF009-002 Manteniment de les dades de la Base de dades de Centres (BDC)	0,00	0,20	0,00	Clúster2
LOP001-001 Desenvolupament del sistema de protecció de dades de caràcter personal	0,91	0,18	0,50	Clúster1
LOP002-001 Atenció al drets d'autodeterminació informativa reconeguts per la normativa de [...]	0,40	0,25	0,25	Clúster1

SÈRIE DOCUMENTAL	ÍNDEX SER-18	ÍNDEX SER-19	ÍNDEX SER-20	CLUSTER
LOP003-001 Registre i notificacions de violacions de seguretat	0,60	0,05	0,17	Clúster1
LOP005-001 Auditories APDCAT i controls de seguretat sobre protecció de dades de caràcter personal	0,67	0,52	0,25	Clúster1
LOP006-001 Manteniment i publicació del registre d'activitats de tractament	0,90	0,20	0,67	Clúster1
LOP007-001 Assessorament en matèria de protecció de dades personals	0,55	0,30	0,43	Clúster1
NEE002-001 Actuacions en casos de sospita de maltractament i/o assetjament en un centre educatiu	0,41	0,28	0,13	Clúster1
NEE020-001 Intervencions en matèria de seguretat ciutadana i prevenció de radicalismes violents	0,47	0,00	0,33	Clúster1
NEE021-001 Seguiment i valoració d'alumnes en situació de vulnerabilitat i/o malestar social [...]	0,31	0,17	0,44	Clúster1
ORG002-001 Gestió de programes en l'àmbit de processos i sistemes	0,75	0,38	0,33	Clúster1
ORG003-001 Anàlisi i millora de processos	0,81	0,31	0,50	Clúster1
ORG004-001 Normalització de formularis i models de documents	0,63	0,37	0,25	Clúster1
ORG004-002 Normalització de formularis i models de document	0,45	0,18	0,25	Clúster1
TRA001-001 Publicació d'informació a l'espai de transparència	0,25	0,25	0,50	Clúster1
TRA002-001 Gestió de les sol·licituds d'accés a documents i dades per part interessada	0,62	0,19	0,33	Clúster1
TRA003-001 Assessorament en matèria de transparència als centres educatius	0,25	0,13	0,20	Clúster2
TRA004-001 Atenció a les peticions d'accés a la informació pública	0,61	0,29	0,33	Clúster1

El resultat, posat en relació amb els tipus de les sèries del grup focal, mostra la següent correspondència:

Taula 7-61: Correspondència entre els clusters de transaccionalitat i els tipus de sèrie documental

CLUSTER	PEL VINCLE FUNCIONAL	PEL TIPUS DE TRAMITACIÓ
Clúster1	Transaccional	Reglada
		No reglada tècnica
Clúster2	Transaccional ²⁸⁶	Reglada
		No reglada tècnica
	Recapitulativa	No reglada temàtica

²⁸⁶ En el cas de la sèrie INF002-001 Eleccions al Consell Escolar de Centre els valors de l'índex de transaccionalitat es poden considerar anòmals, probablement per una entrada de dades incompleta, atès que per la naturalesa de la sèrie hauria de contenir informació com a mínim d'un objecte "acció".

Malgrat estem treballant amb un nombre limitat de dades, observem que l'índex de transaccionalitat, especialment en les modalitats basades en dades i en objectes (SER-18 i SER-20), pot esdevenir un indicador vàlid del caràcter transaccional de la sèrie documental.

Un cop fixada la variable que ens pot determinar el caràcter de la sèrie, i, en l'objectiu 4, la variable que ens determina amb més claredat el caràcter del sistema d'informació, utilitzarem ambdues variables per a fer la comparativa entre sèries i sistemes, i analitzar si existeix una correlació entre els tipus de sèries documentals i els tipus de sistemes d'informació. Aquesta comparativa avalua la correlació entre els tres índex de transaccionalitat de cadascuna de les dues entitats:

- Índex de transaccionalitat a partir de les dades elementals (SER-18 amb SIS-17)
- Índex de transaccionalitat a partir del tipus de dades (SER-19 amb SIS-18)
- Índex de transaccionalitat a partir dels objectes (SER-20 amb SIS-19)

Un cop eliminats els nombrosos valors nuls (existeix un nombre elevat de sèries documentals, i especialment de sistemes d'informació, dels quals no es disposa el detall de les dades elementals), el resultat no ha confirmat una correlació significativa:

Taula 7-62: Correlació entre els índex de transaccionalitats de les sèries documentals i dels sistemes d'informació

	Correlació de Pearson
SER-18 amb SIS-17	0,36
SER-19 amb SIS-18	0,39
SER-20 amb SIS-19	0,47

7.6.1 Conclusió

Un cop analitzades les dades obtingudes, convenim que l'índex de transaccionalitat, i especialment l'índex basat en objectes, que considerem rellevant per classificar una sèrie documental com a transaccional, no determina cap relació entre els tipus de sèrie documental i els tipus de sistema d'informació. Una sèrie documental transaccional, per tant, pot utilitzar sistemes d'informació registrals sense que aquesta utilització afecti el seu caràcter "natural". No s'ha disposat de dades per observar si també es compleix la relació en sentit invers, en el sentit que un sistema d'informació transaccional doni suport o contingui dades d'una sèrie documental recapitulativa.

7.7 Objectiu 6: les visualitzacions de dades com a recurs per a la gestió

La hipòtesi de l'objectiu 6 planteja la possibilitat de considerar les visualitzacions de dades com un tipus d'evidència més, equiparable als tipus ja existents tant en la finalitat com en el valor.

Des de 2019 el Consorci ha desenvolupat 199 visualitzacions de dades, que es distribueixen per la seva finalitat segons les categories següents:

Taula 7-63: Nombre de visualitzacions de dades per tipus de visualització

TIPUS	VISUALITZACIONS
Gestió i estadística	33
Gestió	71
Estadística	95

Si analitzem les 104 visualitzacions vigents a dia d'avui, observem que en un 79% dels casos (82 visualitzacions) són equivalents a una única sèrie documental, tal com s'indica en la taula 7-64. Això ens permet confirmar la possibilitat de considerar la visualització de dades com un document o evidència més de la sèrie documental, i en conseqüència del circuit documental del procés que la genera.

Taula 7-64: Correspondència entre les visualitzacions de dades i les sèries documentals del Consorci d'Educació de Barcelona

VISUALITZACIÓ DE DADES	SÈRIE DOCUMENTAL
QC000_01 Llegenda	
QC001_08 Absentisme_2021	ESC003-002 Seguiment de l'absentisme en el marc de la Comissió de Tractament de l'Absentisme Escolar de Districte
QC001_09 Absentisme_2021_anònim	ESC003-002 Seguiment de l'absentisme en el marc de la Comissió de Tractament de l'Absentisme Escolar de Districte
QC001_11 Absentisme_2223	ESC003-002 Seguiment de l'absentisme en el marc de la Comissió de Tractament de l'Absentisme Escolar de Districte
QC002_03 Ajuts de menjador_2021	BEQ003-200 Ajuts individuals de menjador
QC002_05 Ajuts de menjador_2223	BEQ003-200 Ajuts individuals de menjador
QC004_05 Programes educatius (CUP)_2122	EDU001-200 Programes educatius
QC004_06 QC004_06_Programes educatius i capacitació_2223	EDU001-200 Programes educatius
QC004_07 QC004_07_Programes educatius i capacitació_2223_CRP	EDU001-200 Programes educatius
QC004_08 Programes educatius i capacitació_2324	EDU001-200 Programes educatius
QC004_09 Programes educatius i capacitació_2324_CRP	EDU001-200 Programes educatius
QC005_06 Maltractament_Assetjament_2223	NEE002-001 Actuacions en casos de sospita de maltractament i/o assetjament en un centre educatiu
QC008_04 Fitxa de centre_2122	INF009-001 Manteniment de les dades de la Base de Dades de Centres (BDC)

VISUALITZACIÓ DE DADES	SÈRIE DOCUMENTAL
QC008_05 Fitxa de centre_2223	INF009-001 Manteniment de les dades de la Base de Dades de Centres (BDC)
QC009_05 Consola Gestió Escolarització_2223	ESC005-002 Preinscripció, matrícula i nova admissió
QC009_06 Consola Gestió Escolarització_2324	ESC005-002 Preinscripció, matrícula i nova admissió
QC012_01 BDC_Centres i altres centres	INF009-001 Manteniment de les dades de la Base de Dades de Centres (BDC)
QC015_04 Àrea Inclusiva Expedients NESE_2223	NEE003-002 Reconeixement de l'alumne amb NEE-A
QC017_03 Modificació capacitat centres_2021	PLA010-001 Autorització de modificacions de capacitat
QC020_04 Centres d'Educació Especial_2223	NEE010-001 Sol·licituds de preinscripció, reclamacions i nova admissió en centres d'educació especial (CEE)
QC021_02 Peticions del Banc de Dades	INF001-002 Manteniment del banc de dades del CEB
QC022_01 Base_RALC	
QC024_02 Recursos humans	
QC025_02 Pla d'acompanyament_2021	ESC002-002 Orientació a l'alumnat de 4t d'ESO per estudis post-obligatoris
QC026_04 Postobligatoria_2223	
QC027_02 ECASS	
QC028_02 Indicadors CEB_2021	
QC029_01 TICCONF	
QC031_03 Preinscripció_2223	ESC005-002 Preinscripció, matrícula i nova admissió
QC031_04 Preinscripció_2324	ESC005-002 Preinscripció, matrícula i nova admissió
QC032_03 Alumnat vulnerable_estadístiques_2122	NEE021-001 Seguiment i valoració d'alumnes en situació de vulnerabilitat i/o malestar social o psicosocial
QC034_03 Direccions EAPs_22-23	NEE003-002 Reconeixement de l'alumne amb NEE-A
QC035_02 Anàlisi dades programes innovació (CUP)_2223	EDU001-200 Programes educatius
QC036_01 Seguiment COVID 19_2021_Gestió Interna	ORG008-001 Mesures organitzatives i gestió dels casos covid-19 als centres educatius
QC036_02 Seguiment COVID 19_2021_Estadístiques	ORG008-001 Mesures organitzatives i gestió dels casos covid-19 als centres educatius
QC036_03 Seguiment COVID 19_2122_Estadístiques	ORG008-001 Mesures organitzatives i gestió dels casos covid-19 als centres educatius
QC037_02 Tauletes_2122	
QC038_01 Resolució vulnerables_2021	NEE021-001 Seguiment i valoració d'alumnes en situació de vulnerabilitat i/o malestar social o psicosocial
QC039_01 Gestió de la borsa de docents de suport a centres COVID-19	GRH023-001 Gestió de borses de personal de suport als centres educatius (COVID-19)
QC040_01 PROAVA_2021	
QC041_01 Seguiment_de_projectes_USIDP	ORG002-001 Gestió de programes en l'àmbit de processos i sistemes
QC042_01 Fonts del Banc de Dades del CEB	INF001-002 Manteniment del banc de dades del CEB
QC043_01 Alumnat vulnerable_2021_gestió	NEE021-001 Seguiment i valoració d'alumnes en situació de vulnerabilitat i/o malestar social o psicosocial

VISUALITZACIÓ DE DADES	SÈRIE DOCUMENTAL
QC043_02 Alumnat vulnerable_2122_gestió	NEE021-001 Seguiment i valoració d'alumnes en situació de vulnerabilitat i/o malestar social o psicosocial
QC043_03 Alumnat vulnerable_2223_gestió	NEE021-001 Seguiment i valoració d'alumnes en situació de vulnerabilitat i/o malestar social o psicosocial
QC044_01 Cribratge massiu_SEGUIMENT_2021	ORG008-001 Mesures organitzatives i gestió dels casos covid-19 als centres educatius
QC045_02 Seguiment Preinscripció NEE_2324	ESC005-002 Preinscripció, matrícula i nova admissió
QC045_03 Seguiment Preinscripció NEE_2324_EAPs	ESC005-002 Preinscripció, matrícula i nova admissió
QC046_01 Antecedents documentals dels Serveis Jurídics	JUR003-001 Elaboració d'informes jurídics preceptius
QC046_02 Antecedents documentals dels Serveis Jurídics_2223	JUR003-001 Elaboració d'informes jurídics preceptius
QC047_02 Gestió de la plantilla dels centres públics_2223	GRH024-001 Gestió de plantilles i dotacions dels centres educatius públics
QC048_01 Circuit de resposta a les demandes d'intervenció de les comunitats educatives i	NEE020-001 Intervencions en matèria de seguretat ciutadana i prevenció de radicalismes violents
QC049_01 Orientacions Pla d'acompanyament_2021	ESC002-002 Orientació a l'alumnat de 4t d'ESO per estudis post-obligatoris
QC050_03 Fitxa de centre al mòbil_2223	INF009-001 Manteniment de les dades de la Base de Dades de Centres (BDC)
QC051_02 Activitats estiu_2022	
QC051_03 Activitats Estiu 2023	
QC052_01 Cercador alumnat vulnerable_2021	NEE021-001 Seguiment i valoració d'alumnes en situació de vulnerabilitat i/o malestar social o psicosocial
QC053_02 Renovació juntes directors post-obligatòria 2022	GCE016-206 Juntes de directors
QC054_01 Seguiment signatures AJMENJADOR	BEQ003-200 Ajuts individuals de menjador
QC055_04 Servei menjador_2324	GCE024-201 Gestió dels menjadors escolars
QC056_01 Sistema de gestió documental CEB	DOC001-001 Desenvolupament del sistema de gestió documental
QC058_02 Transport_2223	GCE024-200 Gestió del transport escolar
QC060_01 RALC 0-3	ESC005-002 Preinscripció, matrícula i nova admissió
QC061_02 Seguiment orientacions alumnat 4t ESO_2324	ESC002-002 Orientació a l'alumnat de 4t d'ESO per estudis post-obligatoris
QC062_02 Seguiment alumnat ucraïnès_2223	NEE021-001 Seguiment i valoració d'alumnes en situació de vulnerabilitat i/o malestar social o psicosocial
QC063_01 Resultats Competències Bàsiques	
QC064_01 Seguiment Preinscripció Ensenyaments Postobligatoris_2223	ESC011-003 Preinscripció en ensenyaments post-obligatoris
QC064_02 Seguiment Preinscripció Ensenyaments Postobligatoris_2324	ESC011-003 Preinscripció en ensenyaments post-obligatoris
QC065_01 Anàlisi ensenyaments i concerts_2223	
QC066_01 Ampliem espais	
QC067_01 Gestió de l'alumnat dels Serveis d'Orientació - Pla Jove_2223	ESC002-001 Actuacions dels Serveis d'Orientació
QC068_01 Atenció oficines OME	ATU006-001 Atenció al públic
QC069_01 Seguiment de la despesa	ECO001-001 Gestió de la despesa

VISUALITZACIÓ DE DADES	SÈRIE DOCUMENTAL
QC070_01 Liquidació del pressupost	ECO006-200 Liquidació del pressupost
QC071_1 Consulta Certificats Recursos Humans	GRH020-002 Certificacions idiomes i cursos
QC072_01 Registre de dipòsits de documentació en centres educatius	DOC005-001 Gestió de dipòsits de documentació
QC073_01 Gestió de l'Escolarització	ESC005-002 Preinscripció, matrícula i nova admissió
QC074_01 Registre i marcatge alumnat NEE A	NEE003-002 Reconeixement de l'alumne amb NEE-A
QC075_01 Increments de plantilla_2223	GRH024-001 Gestió de plantilles i dotacions dels centres educatius públics
QC076_01 Renovació juntes directors obligatòria 2022	GCE016-206 Juntes de directors
QC077_01 Ús social centres_2223	INF007-001 Coordinació de l'ús social de centres
QC078_01 UAO	NEE023-001 Gestió de les Unitats d'Acompanyament i Orientació (UAO)
QC079_01 Dades personals docents	
QC080_01 Activitats de vacances_2223	
QC081_01 Educació i Territori_2223	
QC082_01 Fitxa de Centre al mòbil_app	INF009-001 Manteniment de les dades de la Base de Dades de Centres (BDC)
QC084_01 Mapa Adscripcions Centres	GCE010-001 Adscripció de centres
QC085_01 Territori	
QC086_01 Tancament NSDC	
QC087_01 Monitoratge	NEE008-005 Servei de monitors per a alumnat amb discapacitat física escolaritzat en centres d'ensenyament ordinari [...]
QC088_01 Incidències en equipaments	EQU008-001 Manteniment correctiu als centres educatius
QC089_01 Gestió d'inspectors	
QC090_01 Registre Alumnat Inclusiva_2223	NEE003-002 Reconeixement de l'alumne amb NEE-A
QC091_01 Petició de valoració NESE B_2223	NEE021-001 Seguiment i valoració d'alumnes en situació de vulnerabilitat i/o malestar social o psicosocial
QC092_01 Protocols Inclusiva_2223	NEE002-001 Actuacions en casos de sospita de maltractament i/o assetjament en un centre educatiu
QC093_01 Reclamacions PdX_2223.pbix	NEE021-001 Seguiment i valoració d'alumnes en situació de vulnerabilitat i/o malestar social o psicosocial
QC094_01 Càlcul pressupost Pla de Xoc	NEE021-001 Seguiment i valoració d'alumnes en situació de vulnerabilitat i/o malestar social o psicosocial
QC096_01 Càrrega massiva RALC NEE A-B	
QC097_01 Simulacions preinscripció obligatòria_2324	ESC005-002 Preinscripció, matrícula i nova admissió
QC098_01 Seguiment projectes d'obres	EQU002-001 Gestió i seguiment de Projectes de nova construcció, rehabilitacions, ampliacions o construccions modulars
QC099_01 Validació tipus NESE preinscripció postobligatoris_2324	ESC011-002 Validació de NEE en el procés d'admissió d'alumnat als ensenyaments post-obligatoris
QC100_01 Gestió de l'alumnat NEE-A	NEE003-002 Reconeixement de l'alumne amb NEE-A

VISUALITZACIÓ DE DADES	SÈRIE DOCUMENTAL
QC101_01 Quadre de ràtios de centres públics 22-23	
QC102_01 Reclamacions preinscripció_2324	ESC005-002 Preinscripció, matrícula i nova admissió

Però l'anàlisi de les visualitzacions, i especialment del seu ús al Consorci, ha permès constatar un fet que considerem rellevant. El fet és l'important increment en el nombre de dades elementals que es produeix en el diccionari de dades d'una sèrie documental quan s'hi afegeix una visualització com a evidència N del circuit. Si analitzem, a títol d'exemple, l'impacte d'afegir una visualització de dades com evidència N en la sèrie documental NEE002-001 Actuacions en casos de sospita de maltractament i/o assetjament en un centre educatiu, observem el següent:

Taula 7-65: Comparativa entre el diccionari de dades de la sèrie documental NEE002-001 abans i després de la inclusió de la visualització de dades en el circuit evidencial

Diccionari de dades de la sèrie documental NEE002-001 sense visualització de dades	Diccionari de dades de la sèrie documental NEE002-001 incloent visualització de dades
Adreça d'alumne/a	Adreça d'alumne/a
Altres causes	Altres causes
Altres tipus d'agressió	Altres tipus d'agressió
Assetjament en grup	Assetjament en grup
Canals d'assetjament	Canals d'assetjament
Càrrec del docent	Càrrec del docent
Causes	Causes
Codi de centre	Codi de centre
Codi postal d'alumne/a	Codi postal d'alumne/a
Convivents	Convivents
Correu electrònic - treballador director	Correu electrònic - treballador director
Correu electrònic de centre	Correu electrònic de centre
Data naixement de l'alumne/a	Data naixement de l'alumne/a
Descripció dels fets	Descripció dels fets
Entitats actuacions prèvies	Districte de centre
Fets transcorreguts	Entitats actuacions prèvies
IDALU d'alumne/a	Estat actuació UAO
Localitat de residència d'alumne/a	Fets transcorreguts
Motivació de l'assetjament	IDALU d'alumne/a
Nivell educatiu	Localitat de residència d'alumne/a
Nom i cognoms del docent	Motiu assetjament
Nom i cognoms director centre del docent	Motivació de l'assetjament
Nom i cognoms treballador del personal CEB	Nivell de maltractament
Seguiment per part de família	Nivell educatiu

Diccionari de dades de la sèrie documental NEE002-001 sense visualització de dades	Diccionari de dades de la sèrie documental NEE002-001 incloent visualització de dades
Sexe d'alumne/a	Nivell educatiu centre
Telèfon Alumne/a	Nom de centre
Telèfon del director	Nom i cognoms del docent
Tipus agressió	Nom i cognoms director centre del docent
Tipus d'assetjament	Nom i cognoms treballador del personal CEB
Tipus maltractament	Origen del maltractament
Valoració de tancament	Qui fa la comunicació
Zona escolar de centre	Seguiment per part de família
	Sexe d'alumne/a
	Telèfon Alumne/a
	Telèfon del director
	Tipologia d'assetjament
	Tipus agressió
	Tipus d'assetjament
	Tipus maltractament
	Titularitat de centre
	Total alumnes
	Total centres
	Total comunicats
	Valoració de tancament
	Zona escolar de centre

L'increment que es produeix no inclou únicament dades estadístiques creades per la pròpia visualització, sinó també noves dades generades per altres sèries documentals, i incorporades mitjançant la visualització a la sèrie NEE002-001. Podem quantificar aquest increment: un 29% de les dades del nou diccionari de dades les aporta la visualització de dades. El resultat, en tot cas, és que les dades incorporades passen a formar part també de la sèrie NEE002-001.

Però hi ha un segon fenomen derivat d'això: el mapa d'objectes d'aquesta sèrie documental s'incrementa amb un nou objecte, l'objecte "Actuació de UAO". I aquest fet obre la porta a realitzar modificacions en la gestió del procés.²⁸⁷ Les implicacions funcionals del creixement en dades de la sèrie documental són doncs evidents.

Això té un impacte directe en l'índex de ròssec intersèrie (indicador SER-17), el qual s'incrementa quan la sèrie incorpora noves dades al seu diccionari de dades. Si

²⁸⁷ Aquest és un tipus d'actuació de suport psicopedagògic a l'alumnat en situacions socio-familiars complexes o dificultats severes en el rendiment escolar, amb la qual cosa es vincula la causa (el maltractament o assetjament) amb una de les mesures aplicades (el suport UAO).

analitzem l'impacte en el cas anterior, observem un increment de gairebé dos punts en la mitjana de l'índex de ròssec intersèrie:

Taula 7-66: Impacte de la inclusió de la visualització sobre l'índex de ròssec intersèrie

	EVIDÈNCIES	DADES ELEMENTALS	SÈRIES AMB COINCIDÈNCIA	MITJANA ÍNDEX SER-17
Abans d'incloure la visualització	6	32	28	17,47
Després d'incloure la visualització	7	45	28	19,06

7.7.1 Conclusió

La variabilitat constant en les visualitzacions actives, donat el caràcter encara experimental del seu ús com a eina de gestió, no ha permès un buidatge sistemàtic d'informació al nivell de dada elemental que fes possible l'anàlisi a través dels indicadors proposats.

Malgrat aquesta limitació, amb la informació recollida a nivell exploratori ens atrevim a concloure que la inclusió sistemàtica de visualitzacions de dades com a components evidencials d'una sèrie documental modifica tant la matriu de variables com el diccionari de dades de la sèrie documental, i té com a efecte un increment substancial del nombre de dades elementals, del nombre d'objectes, i del potencial d'intercanvi de dades amb altres sèries documentals (índex de ròssec intersèrie). Considerem tanmateix que és una línia de recerca a continuar explorant.

7.8 Objectiu 7: de la classificació funcional a la classificació ontològica

Partirem en l'anàlisi de l'objectiu 7 de la sèrie documental d'un procés monosèrie. El procés triat és NEE002 Actuacions en casos de sospita de maltractament i/o assetjament en un centre educatiu, i la sèrie és NEE002-001, amb el mateix títol que el procés. El diccionari de dades d'aquesta sèrie conté les següents 32 dades elementals, les quals fan referència als següents objectes:

Taula 7-67: Diccionari de dades de la sèrie documental NEE002-001

DADA	OBJECTE	TIPUS
Adreça d'alumne/a	Alumne/a	Cosa
Altres causes	Cas de maltractament o assetjament	Acció
Altres tipus d'agressió	Cas de maltractament o assetjament	Acció
Assetjament en grup	Cas de maltractament o assetjament	Acció
Canals d'assetjament	Cas de maltractament o assetjament	Acció
Càrrec del docent	Docent	Cosa
Causes	Cas de maltractament o assetjament	Acció
Codi de centre	Centre educatiu en general	Cosa
Codi postal d'alumne/a	Alumne/a	Cosa
Convivents	Família / Tutor/a legal	Cosa
Correu electrònic - treballador director	Director/a del centre educatiu	Cosa
Correu electrònic de centre	Centre educatiu en general	Cosa
Data naixement de l'alumne/a	Alumne/a	Cosa
Descripció dels fets	Cas de maltractament o assetjament	Acció
Entitats actuacions prèvies	Entitat	Cosa
Fets transcorreguts	Cas de maltractament o assetjament	Acció
IDALU d'alumne/a	Alumne/a	Cosa
Localitat de residència d'alumne/a	Alumne/a	Cosa

DADA	OBJECTE	TIPUS
Motivació de l'assetjament	Cas de maltractament o assetjament	Acció
Nivell educatiu	Alumne/a	Cosa
Nom i cognoms del docent	Docent	Cosa
Nom i cognoms director centre del docent	Director/a del centre educatiu	Cosa
Nom i cognoms treballador CEB	Treballador/a del CEB	Cosa
Seguiment per part de família	Cas de maltractament o assetjament	Acció
Sexe d'alumne/a	Alumne/a	Cosa
Telèfon Alumne/a	Alumne/a	Cosa
Telèfon del director	Director/a del centre educatiu	Cosa
Tipus agressió	Cas de maltractament o assetjament	Acció
Tipus d'assetjament	Cas de maltractament o assetjament	Acció
Tipus maltractament	Cas de maltractament o assetjament	Acció
Valoració de tancament	Cas de maltractament o assetjament	Acció
Zona escolar de centre	Centre educatiu en general	Cosa

Si agrupem les dades d'aquesta sèrie per objecte, obtenim el següent:

Taula 7-68: Distribució per objecte i tipus de les dades de la sèrie documental NEE002-001

OBJECTE	TIPUS	DADA
Alumne/a	Cosa	Adreça d'alumne/a
		Codi postal d'alumne/a
		Data naixement de l'alumne/a
		IDALU d'alumne/a
		Localitat de residència d'alumne/a
		Nivell educatiu
		Sexe d'alumne/a
		Telèfon Alumne/a
Cas de maltractament o assetjament	Acció	Altres causes
		Altres tipus d'agressió
		Assetjament en grup
		Canals d'assetjament
		Causes
		Descripció dels fets
		Fets transcorreguts

OBJECTE	TIPUS	DADA
		Motivació de l'assetjament
		Seguiment per part de família
		Tipus agressió
		Tipus d'assetjament
		Tipus maltractament
		Valoració de tancament
Centre educatiu en general	Cosa	Codi de centre
		Correu electrònic de centre
		Zona escolar de centre
Director/a del centre educatiu	Cosa	Correu electrònic - treballador director
		Nom i cognoms director centre del docent
		Telèfon del director
Docent	Cosa	Càrrec del docent
		Nom i cognoms del docent
Entitat	Cosa	Entitats actuacions prèvies
Família / Tutor/a legal	Cosa	Convivents
Treballador/a del CEB	Cosa	Nom i cognoms treballador CEB

Aquesta agrupació ens permet donar un tractament discursiu a la sèrie documental, i expressar-la amb la sentència següent:

El/la director/a d'un centre educatiu comunica un cas de maltractament o assetjament, que ha afectat a un alumne i/o docent i ha implicat la participació de famílies i entitats, al Consorci d'Educació de Barcelona.

Si analitzem les sentències parcials de cada evidència, obtenim el següent:

Taula 7-69: Sentències de cada evidència de la sèrie documental NEE002-001

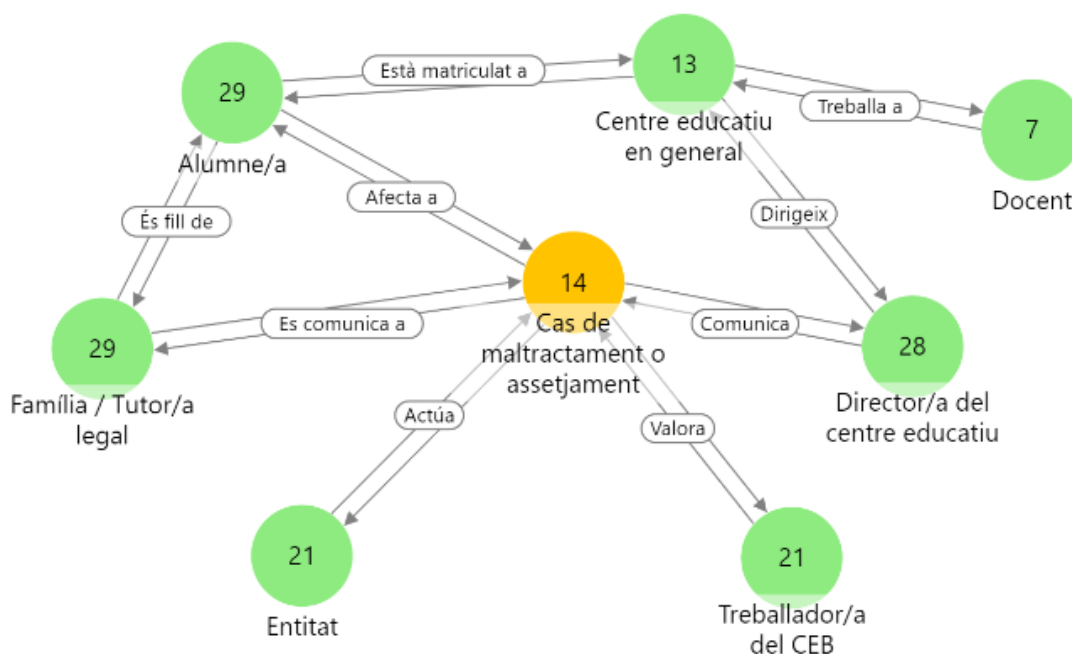
EVIDÈNCIES	SENTÈNCIES
NEE002-E001 Comunicació de cas de maltractament lleu	<i>El/la director/a d'un centre educatiu comunica un cas de maltractament, que pot implicar alumnat, docents i famílies, a un treballador/a del Consorci.</i>
NEE002-E002 Comunicació de cas de maltractament greu	<i>El/la director/a d'un centre educatiu comunica un cas de maltractament, que pot implicar alumnat, docents i famílies, a un treballador/a del Consorci.</i>
NEE002-E003 Comunicació de cas d'assetjament entre iguals	<i>El/la director/a d'un centre educatiu comunica un cas d'assetjament, que pot implicar alumnat, docents i famílies, a un treballador/a del Consorci.</i>
NEE002-E004 Registre i seguiment del cas i activació del protocol	<i>Un treballador/a del Consorci registra i activa protocol pertanyent a cas de maltractament/assetjament.</i>

EVIDÈNCIES	SENTÈNCIES
NEE002-E005 Registre de les actuacions realitzades	<i>Un treballador/a del Consorci registra actuacions pertanyents a cas de maltractament/assetjament realitzades per entitats o docents.</i>
NEE002-E006 Informe de valoració i tancament del cas	<i>Un treballador/a del Consorci elabora un informe valoratiu del cas de maltractament/assetjament, que ha implicat alumnat, docents, famílies i entitats.</i>

Amb aquesta operació obtenim la descripció d'una sèrie documental on apareixen dades de vuit objectes diferents. Un dels objectes és del tipus acció, i és el que articula l'aparició de la resta. S'observa doncs el patró d'una sèrie documental de caràcter transaccional, que queda clarament definida.²⁸⁸

Però desgranar el diccionari de dades de la sèrie entre els objectes descrits obre la porta a una nova possibilitat. Si establim les relacions d'acció entre els vuit objectes que intervenen en la tramitació d'un cas de maltractament o assetjament, incorporem una nova dimensió al tractament de les dades de la sèrie documental. Les dades ja no es relacionen únicament en funció de les evidències on apareixen, del procés que les genera i de l'objecte que descriuen, sinó també en funció de la relació entre els objectes.

Figura 7-28: Mapa de relacions limitades a l'àmbit funcional directe de la sèrie documental NEE002-001



²⁸⁸ Per a la determinació d'aquesta característica vegeu els apartats 7.2 i 7.6 d'aquest capítol.

7.8.1 Conclusió

El cas d'aplicació del Consorci d'Educació de Barcelona ens confirma doncs les opcions 1 i 3 de l'objectiu 7 plantejat a la introducció (apartat 1.3.1.7 del capítol 1), i descarta l'opció 2. Basant-nos en el cas del Consorci, podem afirmar que:

- 1. La classificació funcional és aplicable a un entorn dadificat (i per tant permet classificar sistemes d'informació o conjunts de dades).
- 3. La classificació funcional i els models ontològics són compatibles en un mateix entorn de gestió dadificada.

Aquesta conclusió obre la porta a un replantejament en l'ús i estructura dels sistemes de classificació arxivístics actuals, els quals al Consorci venen donats per estàndards corporatius de les administracions públiques. Aquest replantejament s'exposa i es raona en les conclusions (apartat 8.1.7 del capítol 8).

7.9 Epíleg del cas d'aplicació

El cas d'aplicació del Consorci d'Educació de Barcelona ha permès obtenir un conjunt d'indicadors que esdevenen evidències per a la confirmació o refutació de les hipòtesis plantejades. La informació disponible no ha estat exhaustiva en tot els casos, i per tant hem de considerar aquests resultats vàlids només de forma aproximada, obrint una línia de recerca per, amb un nombre major de dades, poder descartar biaixos propis de la idiosincràsia del Consorci i avançar cap a unes conclusions de validesa general.

És doncs amb aquestes prevencions i en aquest context que es formulen les conclusions que es presenten en el capítol 8.

Capítol 8. Conclusions generals

L'objectiu primer de la tesi és analitzar si el concepte de sèrie documental segueix essent útil, en un entorn basat majoritàriament en dades, per aplicar polítiques de conservació i eliminació sobre els conjunts de dades. Amb una derivada important: si la sèrie com a concepte manté la seva utilitat, també ho fa el criteri funcional que la genera mitjançant la classificació, i en conseqüència les classificacions arxivístiques.

Però la tesi té un segon objectiu, i és posar a prova un mètode de governança de la informació basat en unes eines metodològiques i un conjunt d'indicadors, que són els que s'han utilitzat per avaluar el plantejament anterior.

Com s'ha exposat en l'apartat 6.3 del capítol 6, durant la recollida d'informació no s'ha pogut assolir el mateix grau d'exhaustivitat en la descripció de totes les entitats analitzades, bàsicament pel fet que els mapes han estat en constant evolució, tant fruit dels canvis naturals en l'organització com dels efectes de la transformació digital. Considerem doncs que no disposem de prou dades per assolir resultats significatius mitjançant una anàlisi estadística global. Per aquest motiu, l'estudi s'ha basat sobretot en una anàlisi de caràcter qualitatiu, i només en aquells casos on el volum de dades disponibles ha estat suficient, s'ha aprofundit en l'anàlisi de forma quantitativa.

Les conclusions d'aquest capítol s'estructuren en el mateix ordre que els objectius que s'enumeren en l'apartat 1.3 de la introducció (capítol 1). Per a cada objectiu s'exposen les conclusions detallades, derivades, sempre que ha estat possible, de l'anàlisi del cas d'aplicació, i complementades amb la recerca bibliogràfica. Al final del capítol s'exposa en un epíleg el resum de cada conclusió.

8.1 Objectius teòrics

8.1.1 Objectiu 1: característiques de la sèrie documental dadificada

El cas d'aplicació al Consorci d'Educació de Barcelona ha permès conèixer i analitzar les característiques de les sèries documentals dadificades. Malgrat que el model de dadificació aplicat és un model propi d'un ens en concret, considerem que les conclusions són extrapolables a organitzacions de dimensions i nivell de tecnificació similars, atès que tant l'escenari tecnològic com els tipus de processos són, tal com s'explica en el punt 1 del capítol 6, similars a altres organitzacions del sector públic català.

La primera conclusió derivada de l'observació de les sèries analitzades és que és molt difícil assolir la dadificació completa d'una sèrie documental. Si ens fixem en el primer indicador descriptiu, el nombre d'evidències per sèrie (indicador SER-02 de l'apartat 6.4.2) observem que, fins i tot en el grup focal, no supera de mitjana el 50% de les evidències de les sèries documentals, i que només arriba al 100% en determinats processos d'abast limitat i amb participació d'agents exclusivament interns. Si ens fixem en l'índex de dadificació basat en evidències (indicador SER-13 de l'apartat 6.4.4.2), observem novament que l'índex mitjà amb prou feines supera el 0,5, i que només cinc sèries assoleixen un índex 1 de dadificació completa (vegeu apartat 7.1 del capítol 7).

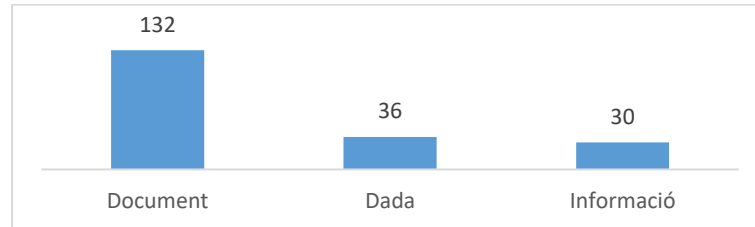
Aquest baix nivell de dadificació es pot explicar bàsicament per tres factors:

- **Factors culturals**, tant de la cultura de treball de la pròpia organització, com de la cultura administrativa-jurídica de l'entorn. Des de la perspectiva interna, moltes organitzacions, especialment del sector públic, tenen encara un imaginari discret de la informació, i conceben l'expedient com una col·lecció de documents (Brooks, 2019). El document encara és referit com a acció, i per tant té una significació pròpia assimilable a una tasca o pas del procediment. El propi llenguatge dels gestors n'és un símptoma, i en la majoria de tramitacions administratives els gestors parlen de "fer la resolució de l'expedient" i no de "resoldre la sol·licitud" o "resoldre el tràmit". En aquest imaginari, els processos es veuen com a col·leccions de documents, i no com a col·leccions de tasques completades (Hoke, 2022; Järvihaavisto & Öhman, 2022).

El context jurídic-administratiu també contribueix a aquesta visió. La *Ley 39/2015, de 1 de octubre, del Procedimiento Administrativo Común de las Administraciones Públicas*, que regula un procediment que, per defecte, estableix com obligatori el

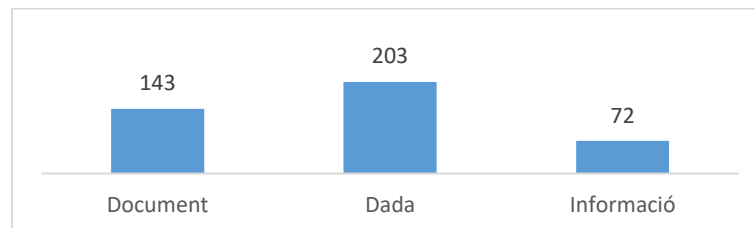
format electrònic tant dels documents (article 26.1)²⁸⁹ com dels expedients (article 70.2)²⁹⁰, mostra un clar desequilibri en l'ús dels termes document i dada:

Figura 8-1: Aparicions del terme dada a la Llei 39/2015



Si ens fixem en una norma més recent, i clarament orientada a promoure una gestió basada en dades, el *Decret 76/2020, de 4 d'agost, d'Administració digital*, observem que la proporció, tot i que s'inverteix, segueix donant un protagonisme important al concepte de document, entès com document-objecte:

Figura 8-2: Aparicions del terme dada al Decret 76/2020



Aquest context jurídic-administratiu configura doncs també l'imaginari social i empresarial que relaciona el sector públic amb la preceptivitat del concepte de document, fins i tot en l'entorn de l'Administració digital.

- **Els documents que s'han de trametre.** Aquest imaginari que hem comentat és compartit amb la ciutadania, que veu en el document un producte del servei que ha rebut (Maroye et al., 2017). Factors generacionals i la introducció dels dispositius mòbils estan canviant ràpidament aquest imaginari, molt més ràpidament en el sector privat que en el sector públic. Probablement l'oferta de

²⁸⁹ "Artículo 26. Emisión de documentos por las Administraciones Públicas. 1. [...] Las Administraciones Públicas emitirán los documentos administrativos por escrito, a través de medios electrónicos, a menos que su naturaleza exija otra forma más adecuada de expresión y constancia."

²⁹⁰ "Artículo 70.2. Los expedientes tendrán formato electrónico y se formarán mediante la agregación ordenada de cuantos documentos, pruebas, dictámenes, informes, acuerdos, notificaciones y demás diligencias deban integrarlos, así como un índice numerado de todos los documentos que contenga cuando se remita."

serveis digitals interactius del sector privat per a dispositius mòbils, superior a la del sector públic, contribueix a accelerar aquest canvi (Van Dijck, 2014). La conseqüència és que, quan es dissenya el procediment, sovint es prioritza la tramesa als agents externs en forma de document, com es pot observar en l'anàlisi de l'apartat 7.2 del capítol 7.

Existeix un segon factor també determinant per a la generació de documents-objecte adreçats als agents d'un procés dadificat. Donar accés a un sistema d'informació propi a agents externs planteja uns reptes importants relatius tant a la seguretat contra intrusions com a la protecció de les dades personals i la informació confidencial. Quan la dadificació es basa en desenvolupaments *low code*, o bé l'entitat no disposa d'un entorn d'interacció autenticada amb agents externs, l'accés per part d'aquests agents pot esdevenir impossible, i per tant caldrà recórrer a la tramesa de documents-objecte.

- **Els mitjans d'autenticació.** L'aposta feta en el seu moment per la signatura electrònica basada en certificat i el desplegament d'una infraestructura pública de certificació electrònica han mantingut també l'imaginari del document com a contenidor exclusiu de referència de la informació autèntica (Alamillo, 2021). L'avenç en aquest cas ha estat important, i el projecte de reglament europeu EIDAS²⁹¹ apunta clarament la tendència cap a la fi de la signatura basada en certificat custodiat pel signant i la introducció d'altres mecanismes d'autenticació aplicables a informació estructurada. La signatura *detached* de conjunts de dades, basada en certificat, i emmagatzemada en camp de base de dades, ja possible anteriorment, es fa ara més fàcil gràcies al reconeixement dels dispositius qualificats remots de creació de signatures (article 3.23 de EIDAS2), és a dir de certificats al núvol que permeten autoritzar la signatura desatesa (Schwalm & Alamillo, 2021). De la mateixa manera que EIDAS2 obre la porta al reconeixement de la validesa legal plena de l'ús de *Blockchain* en el sector públic (Alamillo et al., 2021). Malauradament la regulació a nivell estatal encara manca d'un desenvolupament (Alamillo, 2016, 2019, 2021).

Un indicador molt més significatiu del canvi que aporta la dadificació és l'índex de dadificació a nivell de dada elemental (indicador SER-14 de l'apartat 6.4.4.2). Observem que, malgrat la pervivència d'evidències-objecte en un nombre elevat, algunes sèries ofereixen un índex de dades en format estructurat superior al 0,80. La diferència amb una sèrie no dadificada no està doncs en el nombre d'evidències o documents, sinó en el nombre de dades i, molt especialment, en el nombre d'objectes sobre els quals es

²⁹¹ *Reglamento del Parlamento Europeo y del Consejo por el que se modifica el Reglamento (UE) n.º 910/2014 en lo que respecta al establecimiento de un Marco para una Identidad Digital Europea.*

disposa de dades (indicador SER-08 de l'apartat 6.4.3), cosa la qual representa un potencial d'explotació de la seva informació molt considerable.²⁹²

De fet, és amb l'explotació de les dades que es fan més evidents els avantatges de la dadificació, atès que n'és el principal objectiu (Mordell, 2022). Els processos i les sèries dadificats incrementen la possibilitat de treballar amb el seu contingut de forma estructurada, fent possible construir visualitzacions de gestió on es poden analitzar no només les dades del procés, sinó les de tots aquells processos que, un cop dadificats, incorporen les seves dades al mapa de dades de l'organització. És en aquest punt on una estructura tipus *data lake* o *data fabric* assoleix el seu rendiment òptim. Desenvolupem aquesta conclusió en les conclusions dels objectius 2, 4 i 6.

Ara bé, l'òptim de la dadificació no s'assoleix només amb la disponibilitat del major nombre de dades de la sèrie documental en format estructurat. Un índex alt de dadificació basada en dades (indicador SER-14) no redueix per sí sol la càrrega de treball administratiu, ni contribueix a incrementar la qualitat de les dades, atès que no elimina duplicitats ni garanteix contrastos de validesa. El nivell superior que es pot assolir en una gestió *data driven* s'associa al tradicional ideal de la "dada única", en el qual (Julbe, 2023):

- Es garanteix la qualitat de la dada a través de la seva unicitat i/o sincronització entre tots els sistemes que l'emmagatzemen.
- Redueix la càrrega de treball administratiu, en eliminar la necessitat de tornar a informar dades de les quals l'organització ja disposa.

Els tres indicadors que ens permeten mesurar el nivell d'integració de les dades d'una sèrie documental són l'índex de dadificació ponderat (SER-15), l'índex de ròssec intrasèrie (SER-16) i l'índex de ròssec intersèrie (SER-17).

En el primer cas (SER-15), aquest indicador, que baixa en relació als dos anteriors de dadificació (SER-13 i SER-14) permet corregir biaixos en els dos índex de dadificació anteriors, com per exemple el biaix provocat per l'aparició en una sèrie d'una evidència vinculada a un sistema d'informació de referència que contingui molts més dades que qualsevol de les altres evidències de la sèrie.

La segona mesura és l'índex de ròssec dins de la sèrie documental (SER-16). Aquesta mesura permet objectivar l'ús real de cada dada elemental en el procés, i valorar el seu nivell de reutilització. En aquest cas, observem novament una baixada en relació amb l'índex de dadificació ponderat, cosa la qual ens ajusta encara més les expectatives en relació amb els beneficis de la dadificació de processos. De fet, un índex de ròssec intrasèrie de valor 1,0 pràcticament només el podria assolir un procés que es realitzi de forma íntegrament automatitzada (desatesa). Aquest índex esdevé doncs també un possible índex del nivell d'automatització d'un procés.

²⁹² Vegeu les mètriques de l'objectiu 1 a l'apartat 7.2 del capítol 7.

El darrer índex, el de ròssec entre sèries documentals (SER-17) ja no afecta a una sèrie en concret sinó al conjunt del mapa de sèries de l'organització, i de retruc al mapa de sistemes d'informació. Es tracta d'una mesura d'integració a nivell global. Aquesta integració només pot assolir nivells alts, en temps de gestió, amb un determinat model tecnològic, i especialment amb la implantació de sistemes (centralitzats o sincronitzats) de gestió de dades mestres (MDM). Això no obstant, proposem un segon model tecnològic on, en temps d'explotació, es podran aconseguir nivells d'integració molt alts tot i partint d'un índex de ròssec intersèrie inicial baix. En l'objectiu 2 (apartat 8.1.2 d'aquest capítol) es comenten aquests dos models tecnològics.

Observem en conclusió que la transformació digital basada en solucions *low code*, majoritària en els processos dadificats, permet una ràpida transformació de processos docucèntrics en processos amb nivells considerables de dadificació, i amb un increment molt notable en les possibilitats d'explotació. No obstant, no garanteix una integració real entre els sistemes d'informació que asseguri la fiabilitat de les dades i redueixi el treball administratiu, atès que això depèn de decisions tecnològiques alienes al plantejament documental.

8.1.2 Objectiu 2: característiques de l'ecosistema tecnològic de la sèrie documental dadificada

L'anàlisi de les característiques de les sèries documentals dadificades i dels sistemes d'informació en què es recolzen ens apunta dos possibles escenaris tecnològics on es pot desenvolupar una gestió documental *data driven*.

Les característiques d'una gestió documental orientada a dades són:

- La disponibilitat de totes les evidències i les seves agrupacions, en tant que entitats lògiques, en forma de dades estructurades que puguin ser objecte d'anàlisi i explotació amb els sistemes d'anàlisi de dades actuals.
- La integració de les dades de totes les agrupacions documentals en un model lògic comú, que faciliti la interoperabilitat, tant interna com externa, i permeti la reutilització sistemàtica de les dades fora del seu context original.
- La reducció al màxim possible de la càrrega de treball administratiu relacionat amb la gestió de la informació, i la dedicació dels recursos humans exclusivament a les tasques amb aportació de valor (tecnificació).
- La substitució de les tasques humanes de gestió de la informació per accions absolutament automatitzades.

Aquest resultat es pot assolir, com ha demostrat l'experiència del Consorci, amb dos possibles escenaris tecnològics:

8.1.2.1 Model tradicional

El primer model, que en aquesta tesi hem anomenat model tradicional, consisteix en utilitzar el procés de dadificació com a forma d'identificació dels requeriments funcionals per a desenvolupar una solució tecnològica convencional (Lappin et al., 2021). Com s'ha detallat en la **GDAD**, el procés de dadificació permet construir:

- Una primera definició a alt nivell del model de dades necessari, diferenciant entre les dades transaccionals que hauran de residir en la solució que es desenvolupi, i les dades mestres que s'hauran de gestionar mitjançant un sistema de gestió de dades mestres (MDM), o un sistema d'informació de caràcter registral que en faci funcions equivalents.
- La relació d'interaccions entre agents del procés, indicant en cada cas quines dades s'han de mostrar a cada agent, i quines ha de poder modificar.

- La relació d'evidències-objecte que s'han de poder generar per atendre les necessitats de comunicació amb agents externs que no puguin accedir directament a la solució que es desenvolupi.

Aquesta solució pot consistir en un desenvolupament a mida totalment nou, en l'evolució d'un sistema ja existent, o en l'adquisició d'una solució de mercat. Habitualment serà una solució que utilitzarà bases de dades relacionals, i pot requerir una plataforma de gestió documental com a *back-end* per a la gestió d'objectes digitals. L'automatització de tasques forma part del propi funcionament de la solució, o bé de processos realitzats a nivell de base de dades (Julbe, 2023). El desenvolupament d'interfícies per a la interacció amb agents externs, o la connexió amb portals d'interacció ja existents, formarà part del propi desenvolupament del projecte.

Les polítiques de gestió documental estaran encastades en la pròpia solució, tant pel que fa a la gestió de l'accés com a la conservació i eliminació (generació d'històrics, etc.). En funció de si s'opta per un model *in place* o bé federat amb serveis externs, el sistema pot incorporar funcionalitats de preservació a llarg termini, o bé connectar-se a plataformes externes d'arxiu digital (Lappin, 2010).

Pel que fa a l'explotació, el sistema es pot connectar a qualsevol entorn o tecnologia (*data warehouse*, *data lake*, etc.), assumint la visualització de dades i la creació de *dashboard* com una funcionalitat pròpia o externa.

8.1.2.2 Model *data fabric*

El segon model, que en aquesta tesi hem anomenat model *data fabric*, és un model distribuït (Delgado, 2022; Li et al., 2022; Östberg et al., 2022). En aquest model cada procés disposa d'una solució pròpia, orientada a una gestió bàsica (Moon et al., 2021).

El model tecnològic per al desenvolupament són les plataformes *cloud* de col·laboració, també anomenades de productivitat, els exemples més clars de les quals són els entorns de Google (*Google Workspace*) i Microsoft (*Microsoft 365*). L'ús de tecnologia *SaaS* és preceptiu per aquest entorn, a fi de garantir la connectivitat i la integració amb solucions d'altres processos sense haver d'afegir cap desenvolupament específic (Moon et al., 2021).

El desenvolupament es fa únicament amb parametrització o amb programació lleugera orientada a objectes (*low code*), habitualment a través d'interfícies visuals de programació (Ghiran & Buchmann, 2019; Östberg et al., 2022). Aquesta programació pot implicar la creació de formularis, pantalles d'actualització de dades, aplicacions per a dispositius mòbils, fluxos d'execució programada o a demanda, i robotitzacions de tasques administratives (RPA). Les dades s'emmagatzemen en la pròpia plataforma de col·laboració, habitualment sistemes orientats a objectes que permeten la creació de

sítes, biblioteques, espais compartits, llistes, i que permeten una gestió de les metadades adaptada a cada tipus d'objecte gestionat (Li et al., 2022).

La interacció amb els usuaris, preferentment usuaris interns, es fa concedint accés directament als espais compartits, biblioteques o llistes, a través de formularis, o bé a través dels entorns de productivitat propis d'aquestes suites, com Microsoft TEAMS (Lappin, 2020b, 2021). Tot sovint la interacció amb usuaris externs inclou l'entrada de dades (via formulari), però el retorn es fa per tramesa d'evidències-objecte o enllaços de compartició puntual. En aquest cas no es contempla l'adquisició de solucions de mercat, ni la connexió en temps real amb sistemes d'informació del model tradicional (Delgado, 2022).

Un element especialment important en aquest cas és l'ús de RPA. Davant la impossibilitat de connectar-se directament a sistemes de bases de dades i aplicacions tradicionals, aquestes solucions utilitzen desenvolupaments robotitzats per emular accions d'usuaris humans i intercanviar dades amb aquestes aplicacions externes (Sheperd, 2014).

Les solucions desenvolupades en plataformes de col·laboració són massa senzilles per incorporar polítiques complexes de conservació i eliminació. Així, mentre les polítiques d'accés i la traçabilitat formen part de la pròpia plataforma, la conservació a llarg termini i la preservació digital es delegaran a sistemes externs via flux automatitzat o procés RPA. Permeten per tant una gestió documental *in place* parcial.

La capa d'explotació en aquestes solucions és externa (Ghiran & Buchmann, 2019; Sawadogo & Darmont, 2021; Wieder & Nolte, 2022). Cada aplicació funciona com a font d'un *data lake*, i cal disposar d'una capa de manteniment per a generar el catàleg de dades i els productes d'informació i visualitzacions que siguin necessaris. En aquest sentit, es pot optar per un model basat en la replicació de dades (*data lake* clàssic amb processos ETL/ELT) o per un model basat en la virtualització, que manté les dades en el sistema origen i en garanteix una actualització constant.

Finalment, cal destacar que aquest model facilita la integració amb motors d'intel·ligència artificial, tant pel que fa a la classificació i etiquetat dels continguts com a la identificació i explotació de les dades (Rolan et al., 2019), i probablement incorporarà en breu capacitats generatives (Lappin, 2020c; Patton, 2019).

8.1.2.3 Proposta d'indicadors per a decidir el model

A partir de l'anàlisi qualitativa de l'experiència portada a terme en el cas d'aplicació, observem que es poden identificar, en la tria del model tecnològic, una sèrie de factors objectius, i d'altres d'oportunitat. Relacionem a continuació aquests factors:

Taula 8-1: Proposta d'indicadors per a decidir el model tecnològic per a la dadificació

FACTOR	VALORACIÓ	MODEL TRADIC.	MODEL DATA FABRIC
Factors objectius			
Interacció amb usuaris interns	Si la solució ha d'interactuar únicament amb usuaris que ja tenen accés a l'entorn de productivitat corporatiu, l'opció preferent serà el segon model.		X
Interacció amb usuaris externs	Si la solució ha d'interactuar en algun moment amb usuaris externs, caldrà valorar si la relació és únicament unidireccional, o els usuaris externs han de poder accedir a informació preexistent i modificar-la, cosa la qual requerirà habilitar un sistema d'autenticació. En aquest cas, el primer model serà preferent.	X	
Nivell de seguretat	Si l'aplicació gestiona dades molt sensibles o crítiques per a l'organització, i el nivell de seguretat és elevat, es pot desaconsellar en alguns casos l'ús de solucions SaaS i <i>low code</i> .	X	
Críticitat de la informació	Els entorns del segon model faciliten habitualment un major control dels usuaris finals sobre les dades, però també en fan possible una major modificabilitat. Si els requeriments de consolidació de la informació són elevats, o les dades corresponen a sèries documentals declarades vitals, serà preferent el primer model. El nombre de relacions del procés que genera la sèrie amb altres processos també pot ser un indicador de la críticitat del procés, i del caràcter registral o de dades mestres de l'aplicació a desenvolupar.	X	
Factors d'oportunitat			
Temps de desenvolupament	Les solucions SaaS del model <i>data fabric</i> redueixen de forma molt important el temps de desenvolupament de les solucions, i permeten atendre urgències o solucions de curta durada en el temps. Permeten també posar a prova solucions que, si es consoliden, permeten determinar amb detall els requeriments funcionals d'una futura solució tradicional. Si el temps de desenvolupament és crític, el segon model serà l'opció preferent.		X
Cost de desenvolupament	El cost del desenvolupament ve determinat per la disponibilitat de recursos propis i el temps de desenvolupament. Habitualment les solucions es desenvolupen més ràpidament, i per tant redueixen el cost total del projecte. La vinculació amb sèries documentals prèviament identificades ja facilita un conjunt de requeriments funcionals que redueix el temps d'anàlisi i disseny de l'aplicació.		X
Disponibilitat d'un sistema d'explotació	L'existència d'una infraestructura tecnològica i organitzativa d'explotació de la informació, d'una capa de manteniment i catalogació de les dades, i de capacitat analítica, farà possible que l'aprofitament de la informació sigui el mateix en ambdós casos, potenciant fins i tot la flexibilitat de les solucions del segon model per modificar el seu model de dades i incorporar noves dades amb més agilitat que el model tradicional.		X

Així doncs, davant la necessitat de transformar (i dadificar) un procés, i en funció dels indicadors anteriors, es podrà optar per dues estratègies (Lappin et al., 2021):

- Una estratègia tradicional, més propera al paradigma tecnològic. Es donarà prioritat a l'eficiència del sistema per sobre de la seva significació documental. L'anàlisi funcional tindrà com objectiu identificar el màxim nombre possible de processos als quals podrà donar servei el sistema que es vol desenvolupar, i intentar concentrar en forma de dades mestres el màxim possible d'informació. És el model utilitzat per als sistemes clau d'una organització, els quals habitualment es defineixen a nivell objectual: sistema de gestió acadèmica, centrat en l'alumne/a, sistema de gestió de personal, centrat en el treballador/a, etc.
- Una estratègia *data fabric*. Aquesta estratègia buscarà la correspondència del sistema amb el procés o la sèrie documental, donant prioritat a obtenir un conjunt de dades representatiu d'aquella entitat que després pugui ser objecte d'explotació de forma independent de la solució desenvolupada. Requereix habitualment la disponibilitat de plataformes de desenvolupament lleuger (*low code*) per desenvolupar de forma poc costosa solucions a mida de cada procés, i prou obertes per facilitar després la utilització de les seves dades en un *data lake*. Aquest és el model que més correspondència té amb l'estructura clàssica de sèries documentals, com veiem en les conclusions de l'objectiu 3 d'aquest capítol (apartat 8.1.3). L'anàlisi funcional aportarà en aquest cas dues coses:
 - Un model de dades específic del procés, de caràcter transaccional, que s'implementarà en l'aplicació que doni suport a aquell procés.
 - Una aportació estructurada al model de dades global, que es traduirà en vincles de consum del sistema dedicat amb el sistema de dades mestres, i en l'aportació de nova informació a les dades mestres (creixement orgànic del model de dades global o referencial).

Ambdues estratègies són complementàries, i no exclouen la construcció, quan escaigui, de sistemes robustos per consolidar les dades mestres de l'organització. De fet, la identificació d'un nombre important de dades associades a objectes del tipus "cosa" no vinculades a cap sistema d'informació és un indicador que hauria de promoure en la direcció de l'organització la decisió de construir sistemes que les consolidin.

8.1.2.4 Conseqüències per a la gestió documental

Les conseqüències de l'aplicació d'un o altre model sobre la gestió documental no són trivials, i la validesa dels conceptes i instruments arxivístics pot dependre del model triat. Els dos models es poden combinar en una mateixa organització (i de fet és el que observem en el cas del Consorci), i les característiques, no només del procés, sinó també de la sèrie documental, poden influir en la tria del sistema òptim en cada cas.

El model *data fabric* és un model més flexible, que es pot ajustar a requeriments formals més canviants, i modificar també amb més facilitat quan les característiques d'un procés o sèrie canvien (noves evidències, canvi en els models i formularis, etc.). Aquesta agilitat per adaptar-se als canvis s'ha evidenciat com un element clau en el cas del Consorci, atès que bona part dels processos han incorporat canvis en els seus formularis, models de document, dades de seguiment i necessitats d'explotació pràcticament cada curs, com es pot observar en els valors dels indicadors SIS-02, SIS-03 i SIS-04 de l'apartat 6.6.2 del capítol 6.

El model *data fabric* dona per tant més control al professional de la gestió documental per traslladar els canvis a nivell de sèrie documental a les solucions tecnològiques, fent possible una interacció entre el mapa de sèries i el mapa de sistemes que permet mantenir optimitzats els processos dadificats. És el que James Lappin anomena "*records management by design*" (Lappin, 2021). Això fa possible també que la classificació arxivística, basada en els processos i les sèries, es mantingui sincronitzada amb l'estructura de sistemes d'informació, i permeti traslladar les polítiques de gestió documental a aquests sistemes de forma menys costosa en termes de desenvolupament.

Això tindrà repercussió als efectes d'arxiu, atès que la generació de conjunts de dades per arxivar, com també per publicar en forma de dades obertes, podrà ser definida dins el propi sistema i de forma adaptable als canvis funcionals. El procés de generació de *datasets* d'arxiu, quan es basa en dades en brut (*raw data*) i no prové d'estructures tecnològiques complexes, es facilita pel fet que la coincidència entre el *dataset* i el diccionari de dades de la sèrie documental és gairebé exacta.

En canvi, el model tradicional crea una estructura bàsicament tecnològica, sense *a priori* significació documental, la qual, als efectes d'economia d'escala, tendeix a incloure i transversalitzar funcions de més d'un procés i sèrie documental. Per a ser eficient, el model de dades del sistema ha de ser més complex, i en conseqüència més costós d'adaptar a canvis freqüents. Per aquest motiu la generació de conjunts de dades als efectes d'arxiu (i també de dades obertes) s'haurà de preveure amb antelació, i probablement tendirà a identificar el conjunt de dades amb el sistema en el seu conjunt.

El pes principal de la significació en el model tradicional l'aporta l'estructura del model de dades. Les entitats principals del model, per tant, seran les que orientaran la generació de conjunts de dades d'arxiu, i a vegades fins i tot els processos de gestió. En molts casos això genera una classificació de les dades del sistema marcada per aquesta estructura d'entitats, més afí a matèries o objectes descrits que no pas a funcions i activitats.

Així doncs, a partir de les dades analitzades, tant en el grup focal com en el conjunt de mapes, fem una sèrie d'inferències sobre les característiques arxivístiques d'ambdós models. Plantegem a continuació cinc hipòtesis, i exposem els resultats observats.

8.1.3 Objectiu 3: equivalència entre la sèrie documental i el sistema d'informació

El punt de partida per plantejar-se aquesta equivalència és la conversió de l'expedient en una matriu de dades, tal com s'ha detallat a la **GDAD** (apartat 3.3 del capítol 3).

El coneixement de quines dades dels sistemes d'informació i aplicacions conformen l'evidència d'un procés als efectes d'arxiu és determinant per a l'arxivística, doncs és el que ha de permetre agrupar aquestes dades en sèries documentals per tal de poder-hi aplicar les polítiques de conservació i accés. Així doncs, esdevé cabdal escatir si es poden fer correspondre estructures tecnològiques amb agrupacions documentals, i concretament quines estructures o quines parts d'elles.

Si plantegem l'equivalència al nivell més alt, el de sistema d'informació, podem afirmar el següent:

- La sèrie documental és una entitat que s'identifica a partir d'un criteri funcional i/o documental, i amb un component d'oportunitat (operatiu o pràctic).
- El sistema d'informació és una entitat que s'identifica a partir d'un criteri funcional i/o tecnològic, i amb un component d'oportunitat (operatiu o pràctic).

Per tant, ja identifiquem d'entrada dos possibles punts de contacte:

- El criteri funcional (funció o procés que es vol gestionar). Cada sèrie té un procés que la genera. I cada sistema gestiona tot, una part o més d'un procés.
- El criteri d'oportunitat o operatiu. Cada sèrie respon a una forma concreta de documentar una realitat funcional (forma de crear expedients). I cada sistema respon també a un estil o forma d'automatitzar una realitat funcional, determinada per múltiples factors: pressupostari, tecnològic, de complexitat, de distribució, etc.

I també a nivell de sistema podem valorar un tercer possible punt de contacte: si entenem que el contingut d'una agrupació documental és sempre informació original, el producte directe de l'activitat administrativa o de negoci, podem avaluar l'equivalència d'un magatzem de dades amb l'agrupació documental a partir d'aquesta originalitat.

En aquest punt cal afegir tanmateix un factor de complexitat: la realitat de les organitzacions en un moment donat ve determinada per molts factors (pressupostaris, culturals, etc.), per la qual cosa les característiques de les sèries i dels sistemes sovint no són les òptimes des d'un punt de vista del disseny teòric.

Continuem cercant equivalències analitzant el següent nivell, com s'estructuren les dades dins els sistemes d'informació.

La major part d'aplicacions de gestió utilitzen com a repositori de la informació estructurada **bases de dades relacionals** (Julbe, 2023).²⁹³ Però la forma com la informació s'estructura dins una base de dades relacional no segueix en cap cas criteris documentals, sinó criteris propis del modelat tecnològic (Heuscher et al., 2004; Ramalho, José Carlos et al., 2007). Geoffrey Yeo afirma que els arxivers sempre han intentat cercar la coincidència natural entre les relacions lògiques (el que agrupa evidències en una sèrie) i la implementació física d'aquesta agrupació. I també nega la possibilitat d'aquesta coincidència, argumentant que les relacions lògiques són més complexes del que pugui recollir qualsevol implementació física (Yeo, 2012). Per aquest motiu, és difícil equiparar una taula d'una base de dades relacional amb una sèrie documental.

En el cas dels sistemes d'explotació de la informació tipus **data warehouse**, les estructures relacionals s'han desnormalitzat amb modelat dimensional per permetre una explotació intensiva (vegeu apartat 2.2.3.3 del capítol 2). També és difícil equiparar els **datamarts** a sèries documentals, per dos motius: en primer lloc, perquè són producte d'una reelaboració de les dades originals (no són dades realment originals), que incorporen dades provinents de contextos funcionals diversos, i en segon lloc perquè novament són estructures purament tecnològiques orientades exclusivament a facilitar-ne la interrogació.

En ambdós casos (modelats relacionals i dimensionals) apareix però la possibilitat de generar un objecte digital equivalent a la sèrie documental: la vista o *query*. És a dir, es pot reproduir la instanciació de la matriu de variables d'una o diverses execucions d'un procés.²⁹⁴ Aquesta operació genera al capdavant una taula bidimensional que pot ser objecte de gestió d'una forma similar a com ho pot ser un document-objecte. En aquest cas, la possibilitat d'equivalència amb la sèrie documental (o si escau l'expedient) s'incrementa notablement.

Geoffrey Yeo ho exemplifica dient que (Yeo, 2012):

“Every time I use a search engine on the Internet, I assemble a momentary collection of information about a small subset of that huge collection”.

Aquest raonament és el que ens permet considerar una vista com la manifestació d'una sèrie documental, creada a mida per narrar la història d'uns fets (Bowker & Villamizar, 2017), i posada al mateix temps en context per la seva descomposició entre les entitats d'una ontologia del món gestionat (com comentarem en l'objectiu 7).

²⁹³ En el cas del Consorci d'Educació de Barcelona, la major part dels sistemes desenvolupats pels departaments de tecnologies del Departament d'Educació o per l'Ajuntament de Barcelona (sistemes corporatius) es basen en tecnologia d'emmagatzematge relacional.

²⁹⁴ Vegeu el concepte de matriu de variables a **GDAD**.

Aquest és precisament el procés que s’ha fet a través de l’assignació de metadades a objectes en els gestors documentals (EDMS). En aquest cas, es parteix ja en origen, des del disseny, d’una equivalència dels conjunts d’informació amb entitats documentals, i especialment amb la gestió per expedients, motiu pel qual en el cas d’aplicació la tecnologia que obté un índex més alt d’equivalència entre sèrie documental i sistema d’informació és la orientada a objectes.²⁹⁵

Quan a partir de 2010 es popularitza el concepte de **data lake** com a entorn de treball (Dixon, 2010; Sawadogo & Darmont, 2021; Wieder & Nolte, 2022), el focus d’atenció es comença a desplaçar del sistema cap a la dada en brut. El model *Data lake* té la particularitat de donar cobertura en una mateixa estructura tant a les dades originals (el propi entorn de gestió) com a les dades preparades per a l’explotació. Aquesta cobertura és possible per l’estructuració del *Data lake* en tres capes: la capa d’ingesta, on es dipositen les dades originals (*raw data*), la capa de manteniment, on es consoliden com a dades depurades en una forma similar a les vistes relacionals, i la capa d’exploració, on s’elaboren els productes d’informació necessaris per a l’analítica.²⁹⁶ No és necessari, per tant, estructurar les dades d’una forma determinada, perquè aquesta estructura es pot aplicar en el moment de consumir-les (*schema on read*). Per tant, la dada o el conjunt de dades originals prenen un protagonisme que no tenien en el model relacional.

Baixem al següent nivell, el de dada. Com s’ha exposat a l’estat de la qüestió (apartat 2.2.2 del capítol 2), el concepte de dada, fins i tot quan l’utilitza la ciència de dades, pateix d’una certa imprecisió. El concepte de “dada” és utilitzat per a definir:

1. En primer lloc, la dada atòmica o composta, que es pot entendre com:
 - a. Un valor concret emmagatzemat en un camp de base de dades.
 - b. Una variable o camp de base de dades on s’emmagatzemen valors d’un mateix domini (dada atòmica).
 - c. Un conjunt de variables (o dades atòmiques) que correspon a un concepte identificable com a conjunt en el món real (per exemple nom i cognoms, adreça, etc.). Novament aquest conjunt pot ser un valor concret, o la/les variable/s que pot prendre diferents valors dins un mateix domini.
2. Es pot utilitzar la mateixa definició però sense vincular-la a una implementació concreta en base de dades, amb la qual cosa parlem d’una entitat lògica reconeguda com a tal en un model de dades lògic, o també en un mapa de dades. És el que en el nostre model anomenem “dada elemental” (vegeu l’entitat ENT006 a l’apartat 4.6.1.6 del capítol 4).

²⁹⁵ Vegeu la conclusió de l’objectiu 3 de l’apartat 7.4 del capítol 7.

²⁹⁶ Vegeu l’apartat 2.2.3.4 del capítol 2.

3. Quan el conjunt de variables (o dades atòmiques) va més enllà d'una significació concreta, i adquireix una significació més àmplia, aleshores la “dada” es pot identificar amb un conjunt de conjunts de variables (conjunt de dades compostes), identificable amb el concepte d'entitat o classe en un model de dades (per exemple dada “alumne”, composta de IDALU, nom i cognoms, etc.).²⁹⁷
4. Quan aquesta “dada” (conjunt de dades compostes) és implementada tecnològicament dona origen a components tecnològics concrets, que poden esdevenir objectes de gestió. És quan es parla de “conjunt de dades” (*dataset*). Aquest pot respondre a (Campos, 2023):
 - a. Un fitxer en un format estructurat o semi-estructurat (una taula plana)
 - b. Un model lògic implementat (taula de base de dades, o conjunt de taules i relacions)
 - c. Una base de dades sencera dins un gestor de bases de dades
 - d. Una aplicació informàtica completa que pot utilitzar diferents bases de dades

Retrocedim un moment al nivell anterior. El concepte de “conjunt de dades” independent, entès com objecte de gestió, forma part del nucli de la ciència de dades, i ha penetrat també en la bibliografia arxivística. Així doncs, la majoria de projectes de govern de la dada inicien el seu desenvolupament construint un catàleg de *datasets* (Shanmugam & Seshadri, 2016; Stillerman et al., 2016; Wang, 2014; Zaidi et al., 2017). Però en la majoria de casos no es parteix d'una anàlisi del context funcional i operatiu de cada possible font de dades, ni es parteix d'una lectura holística de l'arquitectura tecnològica de l'organització, sinó que se segueix el procés invers: primer es recullen les fonts de dades, i després s'analitzen per intentar explicar-les i posar-les en context. Una mostra n'és el *business glossary*, el qual, a diferència dels quadres de classificació funcionals actuals, majoritàriament apriorístics, remet a un model anterior, ja existent en l'arxivística, i que es basava en construir el sistema de classificació a mesura que s'anaven localitzant i contextualitzant sèries documentals.

Alguns models de descripció (SIARD, DCAT, etc.) es plantegen descriure *dataset* com a conjunts de dades en abstracte independents de qualsevol implementació tecnològica concreta. Això obre la porta a relacionar la matriu de variables d'un procés, que podem equiparar a la sèrie documental primària d'aquell procés, amb el concepte de *dataset*. El model de metadades **DCAT** (element *dcat:Dataset*) defineix l'entitat *dataset* com:

²⁹⁷ En el nostre model, correspon amb el concepte de “objecte” descrit, que a la seva vegada podem posar en relació amb el concepte “domini” present en un glossari de negoci d'un sistema de governança de la dada. Per a més informació sobre aquest darrer concepte, vegeu l'apartat 2.2.3.5 del capítol 2.

A dataset is a collection of data, published or curated by a single agent. Data comes in many forms including numbers, words, pixels, imagery, sound and other multi-media, and potentially other types, any of which might be collected into a dataset.

Més endavant defineix el concepte de *dataset* com:

[...] collection of data, published or curated by a single agent, and available for access or download in one or more serializations or formats.

Li ho diferencia de les seves diferents representacions tecnològiques, que anomena *distributions*. DCAT especifica que *dataset* és una entitat conceptual, i que pot ser representada per una o més *distributions* que serialitzen el *dataset*. DCAT defineix *distribution* com:

[...] an accessible form of a dataset such as a downloadable file.

Aquesta visió del concepte de *dataset* és hereva del fet que un *data lake* pot contenir qualsevol mena d'informació, sigui aquesta estructurada o no estructurada, amb la condició que sigui processable (Julbe, 2023). Aquest fet contribueix a difuminar la frontera entre el concepte de document, de conjunt de dades i de base de dades, amb la qual cosa dues tradicions de gestió coincideixen sobre el concepte de document: la de la gestió documental, i la de la tecnologia informàtica.

La preservació digital, que ha estat de les primeres actuacions arxivístiques directes sobre les dades, era una oportunitat per introduir aquest raonament i escatir la correspondència entre entitats documentals i tecnològiques (Térmens, 2013; Thomson, 2016; Ur et al., 2015). Malauradament, aquesta transició de la sèrie documental cap a formes tecnològiques basades en dades no ha estat per a l'arxivística una prioritat (Campos, 2023). La preservació d'informació digital estructurada a llarg termini va agafar com a referència inicial el concepte de base de dades en el seu conjunt (Heuscher et al., 2004). Sovint aquest concepte s'ha separat de les aplicacions, considerant com objecte de preservació únicament la base de dades, i no la capa de codi que en facilita la interacció amb l'usuari en un context funcional determinat (Boss & Broussard, 2017).

Però els *datasets*, malgrat es manifestin com objectes independents (fitxers de dades estructurades), gairebé mai es generen directament com a tals.²⁹⁸ Els *datasets* són pràcticament sempre un producte de sistemes d'informació, sovint generats amb una finalitat i voluntat concretes. Qualsevol sistema de base de dades pot generar vistes, tantes com sigui necessari, que tenen un sentit diferent en funció del context d'utilització, sigui aquesta per part d'altres sistemes o per a usos d'explotació. El fet que les dades siguin les mateixes, i provinguin de la mateixa font, no n'unifica els usos, i per tant

²⁹⁸ Poden ser el resultat de la realització d'un tràmit o l'emplenament d'un formulari electrònic, d'un sistema de registre automatitzat amb sensors, o una extracció de dades d'una aplicació, però difícilment es generen originàriament en el format amb el qual ingressen en un *data lake*, llevat dels registres d'activitat que generen els usuaris finals amb eines ofimàtiques.

tampoc el significat. En base a això, ens atrevim a afirmar que els *datasets*, privats del seu context, són un element anti-arxivístic.

Per tant, quan podríem afirmar que un *dataset* és una sèrie documental? Només quan conceptualment coincideix amb el que recull la sèrie documental, és a dir quan té una significació funcional/documental concreta, i no és un mer recull tecnològic o temàtic. En aquest sentit, la matriu de dades d'un procés i l'ontologia que se'n deriva és el que acaba definint la sèrie documental, tal com s'explica en les conclusions de l'objectiu 7.

Un element cabdal en aquesta equivalència entre sèrie documental i conjunt de dades considerem que són les característiques que es defineixen per a les dades mestres (Laurent et al., 2019). En l'objectiu 5 d'aquest capítol apuntem que la sèrie documental, si respon a un criteri funcional, és equiparable a una taula transaccional, i no a una taula mestra. Però des del punt de vista de l'estructuració de la informació d'una organització, el nucli de la informació són les dades mestres, no les transaccionals (Iqbal et al., 2019).

Els tres models que descriu James Lappin en l'evolució dels sistemes de gestió documental són un símptoma de l'evolució que s'està produint (Lappin et al., 2021). El "*Separate records repository model*" desconnecta totalment la responsabilitat dels sistemes origen d'estructurar la informació d'acord amb un criteri arxivístic, i delega al repositori centralitzat la creació de les sèries documentals. No així el "*Intervention in business applications model*", que contempla que les sèries es creïn dins els propis sistemes de gestió, però sense exigir una correspondència sèrie-sistema. És en el tercer model, el model "*in-place*", on observem que, donada la inexistència de funcionalitats de gestió documental en els sistemes origen, la única forma de garantir la creació de context arxivístic (la creació de la sèrie documental) dins el sistema d'informació és forçant la coincidència entre el sistema i la sèrie. Si la relació sistema-sèrie és 1 a 1, la matriu de dades del sistema adquireix la condició de sèrie documental, i les regles es poden aplicar a nivell del conjunt del sistema.

A partir de l'anàlisi de les diferents arquitectures d'emmagatzemament de dades, i de les formes que hi prenen les dades emmagatzemades, proposem en resposta a l'objectiu 3 una relació de les equivalències entre entitats documentals i tecnològiques que considerem possible plantejar, i assenyallem quines s'han pogut establir en el cas d'aplicació del Consorci i quines no, d'acord amb les argumentacions indicades en el capítol 7 i en aquestes conclusions:

Taula 8-2: Possibles correspondències entre entitats documentals i tecnològiques

	Document o expedient	Sèrie documental	Arxiu
Model relacional			
Base de dades	NO	SÍ	NO
Taula	NO	NO	NO
Registre o tupla	SÍ	NO	NO
Vista	SÍ	SÍ	NO
Model data warehouse			
Data warehouse	NO	NO	NO
Datamart i OLAP (cub)	NO	NO	NO
Vista	SÍ	SÍ	NO
Model data lake			
Data lake (i variants)	NO	NO	SÍ
Dataset (sense context)	NO	NO	NO
Dataset (contextualitzat)	NO	SÍ	NO
Model orientat a objectes			
Site	NO	NO	SÍ
Carpeta o biblioteca	NO	SÍ	NO
Llista	NO	SÍ	NO
Objecte o fitxer	SÍ	NO	NO

8.1.4 Objectiu 4: les arquitectures distribuïdes com a entorns de gestió (transaccionals)

La visió més estesa en les organitzacions assigna uns rols funcionals als sistemes d'informació dins les arquitectures de dades en funció del tipus de sistema (Iqbal et al., 2019). Així, les aplicacions sobre bases de dades relacionals es consideren els entorns de gestió per excel·lència, mentre que els entorns d'emmagatzematge distribuït i les visualitzacions d'informació es veuen com elements propis de l'explotació i l'anàlisi de dades, i els entorns de col·laboració com la nova versió de la ofimàtica tradicional (Lappin, 2020).

Però la hipòtesi que hem plantejat en aquesta tesi va en direcció inversa a aquest raonament. Així doncs, fruit de l'experiència d'aplicació en el Consorci, hem plantejat la hipòtesi de considerar els entorns d'emmagatzematge distribuït controlats (el *data fabric*), combinats amb els entorns de col·laboració i les visualitzacions de dades, com els entorns preferents de gestió, i les aplicacions sobre base relacional com els sistemes preferentment utilitzats per a la consolidació de dades mestres.

En l'estat de la qüestió hem definit els diferents tipus de dades que pot contenir un sistema d'informació, i els requisits de la seva gestió (Laurent et al., 2019; Ghavami, 2020). Si tots els sistemes comparteixen una mateixa arquitectura de dades, serà aquesta arquitectura la que distribuirà a nivell de base de dades el tractament dels diferents tipus de dades, i si escau disposarà de sistemes MDM. Però quan no existeixi aquesta integració, el tipus de dades majoritari determinarà la naturalesa funcional del sistema. Això és el que ens permetrà parlar de sistemes d'informació transaccionals, i sistemes d'informació bàsicament registral (gestors de dades mestres).²⁹⁹

A partir del cas del Consorci, s'ha observat que els índex de transaccionalitat de les solucions desenvolupades segons el model *data fabric*, incloent dins aquest model els desenvolupaments fets sobre sistemes orientats a objectes, eren iguals i en algun cas superiors als índex de transaccionalitat de les aplicacions sobre base relacional. Sense que això hagi proporcionat una conclusió definitiva sobre el rol d'una o altra tecnologia, el que ha confirmat és que es poden gestionar processos íntegrament basats en dades només amb tecnologia *low code* en entorns de col·laboració, emmagatzemant les dades en sistemes orientats a objectes, accessibles des d'entorns distribuïts, i que la gestió es pot completar amb visualitzacions interactives de dades, coses les quals faciliten, com hem exposat en la conclusió de l'objectiu 3, l'equivalència entre els sistemes d'informació i les sèries documentals.

²⁹⁹ Una opció propera a l'arxivística és utilitzar el terme "registral" per designar als sistemes que emmagatzemen i gestionen dades mestres.

8.1.5 Objectiu 5: equivalència entre tipus de sèries documentals i tipus de sistemes d'informació

La hipòtesi que es planteja és que els sistemes d'informació que tenen caràcter transaccional són equiparables a sèries documentals (és a dir, es compleix l'equiparació sèrie-sistema), i per tant poden ser objecte, amb les degudes adaptacions, del tractament propi de la gestió documental (conservació/eliminació i accés). Per tant, s'infereix que el concepte de sistema d'informació o aplicació (de gestió o de negoci) comparteix límits conceptuals amb les agrupacions documentals arxivístiques, i té per tant significació documental.

La confirmació d'aquesta hipòtesi obre la porta a que la gestió documental apliqui polítiques directament sobre els sistemes d'informació en el seu conjunt, i consideri les dades de cada sistema com un conjunt amb significació documental (equivalent a sèrie documental).

Per validar aquesta hipòtesi, hem d'assignar prèviament la categoria de transaccional a les sèries documentals, de la mateixa manera com ho hem fet en l'objectiu 4 amb els sistemes d'informació. Però a diferència dels sistemes, hem de partir de la base que, si les sèries s'han classificat (i per tant identificat) d'acord amb un criteri funcional pur, totes les sèries haurien de correspondre a accions; totes serien per tant transaccionals.

No existeix doncs, en un marc de classificació funcional, cap concepte de sèrie documental no-funcional. Però la realitat, que ja hem plantejat en l'apartat 2.3.1.4 del capítol 2, és que:

1. Sovint s'adultera aquest criteri per encabir conjunts de documents que, a nivell operatiu, s'estan gestionant d'acord amb criteris temàtics, objectuals o de conveniència (factuals o facticis). Són les sèries documentals que donen origen als expedients temàtics, també anomenats *subject files*.
2. Les matrius de variables que acaben emmagatzemant determinat tipus de sistemes, especialment aquells que gestionen dades mestres, han de poder-se posar en relació, si és el cas, amb una sèrie documental que faci possible aplicar-hi polítiques de conservació i accés, cosa la qual força a vegades la creació de sèries.

Això no obstant, en aquest objectiu hem acceptat també l'existència de les sèries documentals **recapitulatives**. Això ens permet plantejar el contrapunt a la hipòtesi inicial, de manera que:³⁰⁰

³⁰⁰ De l'anàlisi bibliogràfica de la forma com la tecnologia informàtica identifica quins conjunts de dades són dades de referència, quins són dades mestres i quins són dades transaccionals, es conclou que el nivell de discrecionalitat i, si es pot dir, subjectivitat en aquesta identificació és aproximadament el mateix que trobem en la identificació de sèries documentals. Com en el cas de les sèries, la identificació obeeix sovint a necessitats de tipus pràctic que s'expliquen pel

- A. Les sèries documentals transaccionals s'associaran a sistemes d'informació transaccionals.
- B. Les sèries documentals recapitulatives s'associaran a sistemes de gestió de dades mestres³⁰¹.

Un cop analitzades en el capítol 7 les diferents mesures, concloem que no es detecten diferències significatives en les característiques de les sèries, ni es pot establir un patró diferencial, cosa la qual ens porta a convenir que, en les sèries dadificades, el concepte de sèrie recapitulativa es difumina fins gairebé deixar de tenir sentit. Aquest concepte té la seva base en formes documentals antigues, amb l'objectiu de diferenciar els formats registrals (llibres de registre) dels formats documentals (col·leccions d'objectes). En el moment en què una sèrie es dadifica, tota sèrie es converteix en un registre. Qualsevol registre informàtic recapitula informació, la condensa, trencant la divisió entre documents en tant que unitats documentals simples.³⁰²

Un cop determinat que totes les sèries documentals, si s'aplica un criteri funcional, són transaccionals, s'ha avaluat la correspondència sèrie documental – sistema d'informació.

El resultat no ha estat concloent. La correspondència sèrie - sistema és una correspondència difusa, que depèn de la tecnologia triada, però sobretot de decisions preses en la fase de disseny. Això li confereix un caràcter facultatiu, que fa que aquesta correspondència esdevingui al capdavant una acció voluntària dels equips de disseny, tant dels sistemes d'informació com dels mapes de sèries documentals, i que qualsevol organització pugui ajustar, si així ho desitja, la gestió de les seves dades en sistemes d'informació a criteris i mètodes propis de la gestió documental. Certes arquitectures de sistemes, especialment les arquitectures distribuïdes, fan més fàcil aquest disseny, però no el predeterminen si no existeix la voluntat d'orientar la gestió de les dades a la gestió d'agrupacions arxivísticament significatives.³⁰³

context concret on es realitza aquesta identificació, però que poques vegades són exportables a nivell general, més enllà d'allò que ja ve fixat pel costum (per exemple el registre de clients) o la normativa tècnica (per exemple els codis de països).

³⁰¹ Aquesta correspondència és la que es posa a prova en aquestes conclusions amb diferents indicadors.

³⁰² Utilitzant terminologia ISAD-G.

³⁰³ Això no exclou que internament sigui adequat articular sistemes de gestió de dades mestres (MDM) que consolidin dades, però no necessàriament amb un model centralitzat.

8.1.6 Objectiu 6: les visualitzacions de dades com a recurs per a la gestió

Ja el 2013 Terry Cook, quan parlava de les quatre dimensions de l'arxivística, plantejava l'evolució de l'arxiu des d'un mer custodi de la veritat (autenticitat) a algú que ha d'explicar la història i facilitar els mitjans per entendre-la (Cook, 2013). El canvi en relació amb la prestació de serveis arxivístics en fons que contenen dades no és nou, i el pas des d'una prestació passiva cap a la generació de productes de valor afegit ja ha estat proposat i practicat en nombrosos casos en el camp dels arxius històrics (Yakel, 2000). La necessitat de començar a vincular la gestió documental amb la visualització de dades ja era apuntada indirectament l'any 2005 per un equip de la Universitat de *British Columbia* (Cloonan et al., 2005), en el marc del projecte InterPARES 2, quan parlaven del tractament dels documents experiencials, interactius i dinàmics. I més recentment, des de la mateixa Universitat s'ha reiterat la proposta, ja directament vinculada a les visualitzacions de dades (Lemieux et al., 2014).

L'arxivística administrativa no ha estat tanmateix especialment activa en potenciar la difusió dels fons d'arxiu. A les organitzacions, tant públiques com privades, la transformació de la informació en productes auto-explicables i llestos per al consum immediat ha estat portada a terme majoritàriament per gabinets d'estudis i estadística, i darrerament per les oficines de dades.

Però més enllà del producte analític o estadístic, habitualment generat a demanda o de forma prospectiva, i adreçat especialment als nivells estratègics de l'organització, sempre ha existit una demanda important de producte d'informació operatiu, adreçat a resoldre les necessitats de la gestió diària i la presa de decisions a curt termini dins l'operativa dels procediments. Aquesta necessitat s'ha cobert habitualment per dues vies:

- Els registres d'autoconsum elaborats amb eines ofimàtiques (fulls de càlcul de seguiment i anàlisi, agendes, etc.).
- Les funcionalitats recapitulatives i estadístiques de les pròpies aplicacions informàtiques (l·listats, mètriques i visualitzacions integrades en la pròpia solució).

L'increment de la capacitat tecnològica per al processament i l'analítica de dades, i la consumerització de les solucions de tractament, visualització i anàlisi, juntament amb l'extensió dels models descentralitzats d'emmagatzematge i gestió de dades, han posat a l'abast de l'usuari final la generació de productes d'informació molt sofisticats (Pérez-Montoro, 2020). I aquests productes, en una cultura eminentment visual com l'actual, es basen cada cop més en visualitzacions gràfiques i interactives de dades (Joseph & Hartel, 2017; Lemieux, 2015; Lemieux et al., 2014).

Aquest fenomen ha tingut una derivada, que ha estat la simplificació en el desenvolupament d'aplicacions, pel fet que tota la capa de funcionalitats dedicada a tractament i visualització de les dades s'ha pogut segregar del desenvolupament i realitzar fora de l'aplicació i de forma independent d'aquesta. Per tant, existeix la possibilitat de generar productes d'informació de valor afegit dins l'execució d'un procés sense necessitat de desenvolupar una aplicació informàtica o modificar-ne una d'existent. Els models d'emmagatzemament basats en arquitectures distribuïdes tipus data lake, on s'ha segregat la capa d'explotació de la capa de gestió, han contribuït a fer-ho possible (Reina, 2023).

Però hi ha una segona derivada amb major impacte sobre la gestió documental. Tradicionalment els informes, memòries i estadístiques ha estat la forma (documental) amb la qual s'ha introduït en un expedient o sèrie documental l'anàlisi d'un procés. En altres casos, la inclusió de llistats, o el manteniment de registres de seguiment i control, també ha format part de les solucions per fer constar dins el substrat evidencial d'un procés les dades de suport a la gestió. Des de fa temps, aquests productes es materialitzen en forma de visualitzacions interactives, agrupades en quadres de comandament (*dashboards*), que cada cop poden ser tractats com objectes més independents.

S'obre la porta doncs a considerar el quadre de comandament d'un procés com el document o evidència N del circuit documental del procés, considerant la visualització de dades com una nova forma documental. I aquesta afirmació es reforça amb el següent argument: des del moment en què la visualització de dades s'utilitza per a prendre decisions, que produeixen efectes factuais, es converteix en una evidència de ple dret, que en conseqüència cal arxivar. Aquesta estratègia s'ha posat en pràctica en el projecte del Consorci d'Educació de Barcelona, i ha aportat algunes constatacions, que argumentem quantitativament a l'apartat 7.8 del capítol 7.

La primera és la correspondència de la visualització de dades amb la sèrie documental. Els dos tipus de sèrie documental que han generat visualitzacions dedicades han estat:

- Les sèries reglades, amb un alt volum de tràmit força estandarditzat.
- Les sèries de coordinació d'àmbits, i les que requereixen un alt consum de dades mestres.

Això no obsta per a que també s'hagin generat visualitzacions purament estadístiques o recapitulatives, opcionalment assimilables a sèries vinculades a la funció de planificació estratègica, o a grups de sèries relacionades.

El gran benefici que aporten les visualitzacions a un procés prové precisament d'aquest increment de dades, doncs, fruit de la integració de fonts diverses en què es fonamenten les visualitzacions, es posa a l'abast de l'usuari gestor un nombre important de dades de context, provinents d'altres processos, que l'ajuden a orientar i millorar la seva presa de decisions, fomentant la intel·ligència col·lectiva (Kudyba et al., 2020).

I la tercera derivada és la constatació que la introducció de les visualitzacions com una evidència dels processos té un efecte incremental, que beneficia no només al procés on s'introdueix, sinó també a la resta de processos del mapa global. I, per tant, un impacte important en la millora funcional de l'organització. Les visualitzacions no són per si mateixes un element d'integració, però requereixen la integració prèvia de les fonts de dades per a poder-se generar. Així doncs, cada cop que s'elabora una visualització a partir de les dades d'un procés dadificat, s'obre una via de reutilització de les dades d'aquest procés en la resta de processos de l'organització, atès que es converteix en una nova font potencial de la resta de visualitzacions. D'aquesta manera es contribueix a transversalitzar la disponibilitat i ús de les dades.

Aquest, que és un indicador d'integració de la informació d'una organització, demostra per tant que, amb independència de la reutilització en tramitació de les dades, la seva mera visualització ja esdevé una forma d'integració. Considerem doncs aquesta una línia de treball que caldrà explorar.

8.1.7 Objectiu 7: de la classificació funcional a la classificació ontològica

En el desenvolupament del projecte, per a la identificació de les entitats documentals vam partir del model hegemònic, a part que normatiu: el model funcional. Qualsevol agrupació de documents, doncs, obeeix a un criteri funcional, que en el nostre cas s'estructura per processos relacionats i les sèries que se'n deriven. Aquest model determina les relacions entre sèries, i esdevé doncs el nivell al qual es planteja l'aplicació de polítiques de conservació i accés.

Durant el procés de dadificació dels processos, i en conseqüència de les sèries documentals, vam observar que un dels principals reptes és identificar la dada.³⁰⁴ Atès que el nom no és suficientment descriptiu de la dada, es va observar que la identificació de la dada, per a ser prou precisa, requereix la combinació de dos elements: el nom de la propietat descrita, i el nom de l'entitat a la qual pertany aquesta propietat.³⁰⁵ Això correspon al que en els *business glossary* d'un sistema de governança de la dada s'anomena el domini de la dada (Sarkar, 2015). Evidentment, amb això no s'elimina la indefinició en la identificació, però es redueix notablement. Es va observar també que això ens apropava a la lògica de predicats utilitzada per a la construcció d'ontologies, revelant en el cas de les propietats únicament relacions de comprensió (Feilmayr & Wöß,

³⁰⁴ Vegeu les dificultats d'aquesta tasca en la introducció del capítol 6.

³⁰⁵ De la propietat també en podríem dir atribut. En el nostre projecte, l'entitat descrita l'hem anomenada "objecte", i la propietat o atribut que descriu aquesta entitat l'hem anomenada "dada elemental". Per a més detall vegeu les entitats **ENT006** i **ENT014** del capítol 4.

2016). Així doncs, es van introduir en el projecte alguns elements d'aquesta lògica de predicats, considerant les evidències, i en alguns casos també els sistemes d'informació, com l'expressió de sentències predicatives.

En la lògica de primer ordre, els predicats són tractats com a funcions que reben un conjunt de coses, les processen, i tornen com a resultat una única cosa (Quílez, 2011). Vam aplicar doncs aquesta conceptualització a allò que havia de ser en primer lloc objecte de descripció: els processos, incloent les seves diferents manifestacions (funcions, activitats, transaccions, tasques, etc.). Podem trobar un paral·lelisme evident entre la definició de procés i una funció matemàtica, on existeixen unes entrades, unes sortides i un processament (Serifi et al., 2013; Mokhtar et al., 2016). És un esquema que es pot aplicar tant a nivell de procés, com de sèrie, i fins i tot de tasca o acció concreta, que és el que queda reflectit en una evidència (documental).

Aquest patró d'identificació es va recollir a les dues guies metodològiques, la **GDAD** i la **GMAP**, i es va aplicar al buidatge realitzat per construir el mapa de dades. El buidatge va donar com a resultat una relació de dades elementals, identificades per un nom i un objecte descrit, però no un mapa.

Les dades elementals s'han concebut en aquest projecte com entitats atòmiques, unitats mínimes d'informació significativa. De forma natural, doncs, no s'han establert relacions entre dades. Tota relació entre les dades s'obté doncs del context d'aquestes dades, un context que té dues dimensions:

- Context d'utilització: en quines activitats s'utilitzen (processos, sèries, evidències, sistemes).
- Context de pertinença: a quins objectes descriuen, i com es relacionen aquests objectes entre ells.

Per tant, les dades heretaran les relacions que mantinguin els processos, sèries, evidències i sistemes on apareixen, i també les que mantinguin els objectes descrits entre ells.

Com hem comentat, aquest plantejament ontològic ja era present en el concepte de sèrie documental proposat per Peter Scott (Scott, 1966), que contemplava el marc complet d'interaccions entre agents el qual, al capdavant, acabava construint la sèrie documental, i que va donar lloc a l'*Australian series system* (Hurley, 2008). La proposta d'utilitzar plantejaments ontològics per a la descripció de documentació administrativa no és nova (Yeo, 2013), i es troba clarament desenvolupada en els treballs de Klischewski (Klischewski, 2003, 2006) i d'Alalwan i Thomas (Alalwan & Thomas, 2011, 2012). Ens remetem doncs a totes les experiències citades en l'estat de la qüestió, algunes de gran abast, com la de l'Administració de Portugal (Lourenço et al., 2020; Penteado, 2015, 2018), i també de l'àmbit estatal (Delgado, 2010; Casellas, 2014; Quílez, 2017; Aguayo, 2021). De fet, podem afirmar que el desenvolupament d'ontologies és una de les línies de recerca més actives actualment en la professió

arxivística, sobretot pel seu potencial d'utilització per a l'entrenament d'algoritmes d'intel·ligència artificial (Beheshti et al., 2020; Frontoni et al., 2022; Rolan et al., 2019).

Però la proposta que fem no es planteja la part operativa o d'implementació d'aquest plantejament, sinó la forma de traslladar-hi, adaptar-hi o, si escau, superar el criteri funcional en les classificacions arxivístiques. Per aquest motiu, hem analitzat diferents casuístiques per veure fins a quin punt el pas d'un mapa de processos a un plantejament ontològic conserva el criteri funcional, l'enriqueix, o fins i tot el substitueix. No ens hem plantejat en cap moment com objectiu del projecte la implementació física de l'ontologia en RDF o OWL, només la seva formulació lògica als efectes de la comparativa que proposem. El model ontològic que hem utilitzat d'inspiració ha estat PROV-O,³⁰⁶ incorporant conceptes propis de la desnormalització de bases de dades segons el model dimensional (Ur et al., 2015; Julbe, 2023).

El vincle amb la conceptualització de PROV-O es fa més evident quan analitzem com s'estructuren les entitats descrites en el desenvolupament d'un sistema d'informació. En descriure com s'identifiquen, durant el modelat de dades, quines dades són dades mestres i quines transaccionals, Ghavami afirma que aquesta relació pot ser vista com una relació entre substantiu i verb (Ghavami, 2020). Les dades transaccionals descriuen verbs (és a dir accions), mentre que les dades mestres descriuen substantius (és a dir coses). Aquesta mecànica coincideix amb la que s'utilitza quan s'elabora un quadre de classificació funcional o un mapa de processos, en el qual l'anàlisi s'adreça a identificar en primer lloc els verbs, les accions, els quals després es posen en relació amb les coses afectades per les accions. I és la mecànica que vam utilitzar en aquest projecte quan vam començar a identificar dades elementals, atribuïnt a cada dada el substantiu o verb al qual descrivia, el que després es va convertir en objectes tipus "cosa" i objectes tipus "acció". De la mateixa manera, també vam identificar com a mínim un nivell d'agrupació en els objectes, corresponent a l'agrupació de les dades mestres en dominis i subdominis a la qual fa referència Ghavami (Ghavami, 2020).

El procés, i la sèrie documental, es defineixen doncs com la manifestació discursiva de l'ontologia als efectes de fixar un procediment. Però les evidències d'informació que generen no s'han d'agrupar necessàriament en funció del procediment seguit, sinó que es poden classificar per l'objecte que realment descriuen.³⁰⁷ El procés no té doncs necessàriament una significació temàtica o objectual, sinó que manté una finalitat purament operativa, vinculada al procediment. Aquesta connexió considerem que és possible i beneficiosa, en primer lloc, perquè la definició de procés serveix per derivar el circuit documental i els vincles amb sistemes i agents, i en segon lloc perquè tota part temàtica/objectual es pot obtenir si es classifiquen les dades per objectes. Serà a partir

³⁰⁶ Vegeu <https://www.w3.org/TR/prov-o/>.

³⁰⁷ La gestió de metadades en un *data lake* implica gestionar dades i objectes a diferents nivells o granularitats (Chen 2022).

de l'ontologia que es recuperarà la informació agrupada per objectes descrits, no pas a partir del procés.

En aquest context, el concepte de sèrie documental es pot assimilar al concepte d'entitat de tipus "acció".³⁰⁸ No s'esdevé així amb les entitats de tipus "cosa", que no responen a un criteri funcional, i que en tot cas s'haurien de relacionar amb el que tradicionalment s'han considerat sèries documentals recapitulatives, temàtiques, o no "naturals" (vegeu al respecte l'apartat 2.3.1.4.1 del capítol 2).

Si ens quedem en aquest punt, el concepte de sèrie documental només s'ajusta parcialment a la realitat d'un entorn dadificat. Això no obstant, per tal de treure profit de la utilitat del concepte en la doctrina arxivística (models de descripció ja existents, sistemes de regles de conservació i eliminació, etc.), se'n pot plantejar una reinvençió contemplant dos possibles tipus de sèrie: la sèrie "transaccional" (equivalent a objectes de tipus "acció") i la sèrie "referencial" o "registral" (equivalent a objectes de tipus "cosa"). Podem d'aquesta manera reconciliar el concepte de sèrie documental amb els components d'un model de dades, que sí que encaixen perfectament amb aquesta conceptualització. Així doncs, les entitats i/o taules transaccionals d'un model de dades seran equivalents a objectes del tipus "acció", mentre que les taules o conjunts de dades mestres seran equivalents a objectes del tipus "cosa".³⁰⁹ Cal tenir present que l'associació no es fa al nivell de sistema d'informació o aplicació, el qual és una construcció purament tecnològica i adreçada a donar suport a un conjunt de procediments o seqüències d'activitats, però sense una significació documental.

Així doncs concloem que **la forma òptima de classificació del substrat informacional d'una organització, quan aquest està basat en evidències dadificades, és mitjançant una ontologia administrativa que relacioni coses i accions, implementada, si es disposa d'un entorn de governança de les dades, en forma de *business glossary*.** Això no exclou la utilització d'un mapa de processos per identificar les seqüències d'activitats en fase de gestió i derivar-ne els circuits evidencials, però la recuperació de la informació, l'aplicació de polítiques de conservació i accés, i la generació de versions d'arxiu dels conjunts de dades es considera que serà òptima quan s'obtingui de l'ontologia administrativa i tingui en compte les relacions que s'hi estableixen. La sèrie documental adquireix d'aquesta manera una perspectiva relacional, que reforça el caràcter interseccional ja expressat en la norma ISAD-G, i comentat en l'estat de la qüestió. Així doncs, les dades poden descriure activitats (o processos) en tant que entitats transaccionals expressades mitjançant verbs

³⁰⁸ Vegeu les equivalències amb els diferents models de descripció descrites en el capítol 4, on es detalla el model de dades per a les entitats procés, objecte i dada elemental.

³⁰⁹ El principal risc, com en tota classificació, són els biaixos o la càrrega ideològica. Tota classificació amaga un discurs. Un discurs entès no només com a producte d'una determinada visió del món, sinó també com la voluntat de produir o provocar efectes. Per tant, tota classificació implica una intencionalitat (Mordell, 2019, 2022).

(habitualment substantivats), i també poden descriure coses del món real, sovint agents, però també productes de l'activitat dels agents, en tant que entitats de referència expressades mitjançant substantius. I en aquest context, el que posa en relació, des de la perspectiva de la gestió de la informació, les accions i les coses (agents i productes) és la sèrie documental. En aquest context, **podem definir la sèrie documental dadificada com el conjunt de relacions de l'ontologia administrativa que es poden instanciar en un moment i lloc concrets**, identificant d'aquesta manera el qui, el què, el com, el quan i l'on de cada acció.³¹⁰

Una derivada d'aquesta proposta de treball a dos nivells inclou l'afectació sobre les relacions entre sèries documentals. Així, podem inferir, a partir dels valors de l'indicador SER-17 que es descriuen en l'apartat 6.4.4.3 del capítol 6, que determinades relacions entre sèries documentals es desenvolupen en la capa funcional, mentre que d'altres ho fan en la capa objectual o ontològica. En la primera capa trobarem les relacions de continuïtat (apartat 2.3.1.4.2.2), de derivació (apartat 2.3.1.4.2.3) i de dependència (apartat 2.3.1.4.2.4), mentre que en la segona capa trobarem les relacions estructurals (apartat 2.3.1.4.2.1) i especialment les de contingut (apartat 2.3.1.4.2.5).

L'ús d'ontologies ens obre la porta a representar la sèrie mitjançant grafs, podent fer una lectura focal o sectorial, o bé una lectura global del lloc que ocupa la sèrie dins el conjunt d'informació de l'organització, entesa com un contínuum de dades interrelacionades (Chessa et al., 2022; Schmid et al., 2019; Smajevic et al., 2021; Smajevic & Bork, 2021).

Aquesta és la visió que representem en la figura 8-3. El model *data fabric* ens ofereix la possibilitat d'implantar aquesta concepció de sèrie documental utilitzant els nivells en què s'estructura. I sobre aquesta base ens atrevim a apuntar una proposta conceptual: **assimilar sèrie documental a conjunt de dades, amb una significació diferent en funció de la capa del sistema d'emmagatzematge on ens trobem**. Així doncs:

³¹⁰ Aquestes cinc preguntes corresponen als cinc elements mínims amb els quals podem descriure un expedient: el qui (productor), el què (objectiu o resultat), el com (acció o procés), el quan (dates cronològiques extremes) i l'on (datació tòpica). El vincle amb l'arxivística és doncs directe.

Les noves eines per a la gestió de catàlegs de dades ja contempen aquesta orientació en la forma d'estructurar el *Business glossary*. De fet, una de les tasques cabdals de la integració de dades (en un *data lake*) és la *entity resolution*, que consisteix en encaixar (*matching*) diferents representacions d'entitats provinents de fonts diverses. Aquest procés utilitza tècniques de *blocking* (construcció de blocs per entitats a partir de l'anàlisi de conglomerats d'entitats similars) i de *meta-blocking* (reestructuració dels blocs anteriors en nous blocs per eliminar redundàncies).³¹² I l'assignació de *topic terms* a continguts digitals utilitzant intel·ligència artificial va també en la mateixa línia (Diamantini et al., 2021).

Aquesta proposta no s'ha d'entendre com una negació del criteri funcional, però sí com una relectura on els agents i objectes de les activitats se segreguen de l'essència pròpia de cada activitat.

Els principals avantatges d'implementar aquest model de classificació són els següents:

- L'estructura de classificació és molt propera a com s'estructuraran les dades, especialment en arquitectures *data lake* o *data fabric*, però també en models relacionals clàssics i models amb MDM. La inclusió de funcionalitats per a la generació de versions d'arxiu es fa doncs més fàcil i ràpida.
- Un procés docucentric duplica informació, i especialment dades. Un dels principals avantatges d'un model dadificat és que facilita apropar-se al model de dada única. Però mantenir el criteri d'agrupació a nivell de procés i sèrie documental perpetua la duplicació d'informació, atès que la mateixa dada formarà part de diferents sèries documentals. Per aquest motiu es considera el model d'ontologia d'objectes com més eficaç per evitar les duplicitats en el manteniment i generació de dades d'arxiu. De fet, la dada única s'assoleix a nivell d'objecte, no de procés.
- L'aplicació de polítiques d'accés serà més precisa si s'aplica al nivell d'objecte. Les dades personals o sensibles habitualment es concentren en les entitats "cosa", i l'anonimització sovint actua en aquest sentit. El control de l'accés pot ser més efectiu amb aquesta orientació.
- El mateix passarà amb les regles de conservació. Habitualment, les regles de conservació determinen excepcions en la resolució per tal de preservar les dades mestres i eliminar la informació transaccional. Orientar la classificació a objectes permetria evitar aquestes excepcions, molt sovint de complexa aplicació.

³¹² Aquesta tècnica (anomenada *query driven entity resolution*) ha estat especialment estudiada per l'equip de George Papadakis (Papadakis, Papastefanatos, i Koutrika 2014; Papadakis et al. 2015; Efthymiou et al. 2017).

8.2 Objectius instrumentals

Finalitzem aquestes conclusions amb una valoració general de la metodologia utilitzada per a la recollida i anàlisi de les dades, la qual al mateix temps es proposa per iniciar una organització en la governança de la informació orientada a dades. Els aprenentatges que aquí es comparteixen provenen de l'experiència portada a terme al Consorci d'Educació de Barcelona, i també d'algunes experiències anteriors que se citen en el punt 3.1.2 del capítol 3.

8.2.1 Punts febles

El primer punt feble a destacar és el cost de recollida i manteniment de la informació. Es considera pràcticament impossible mantenir actualitzat el sistema de governança sense tres condicions:

- La primera és integrar la recollida d'informació sobre cada ítem del sistema en operacions de gestió d'execució habitual. El manteniment del sistema de governança no pot ser una tasca afegida als equips de treball. Així doncs, per exemple, si el sistema de governança s'utilitza per al seguiment dels desenvolupaments tecnològics, serà molt més fàcil mantenir actualitzat el mapa de sistemes, i seran més precises les seves dades.
- La segona és l'entrenament dels equips de treball. Tots els professionals que analitzen processos, identifiquen sèries, desenvolupen sistemes, etc. han de rebre un entrenament específic, que cal revisar i actualitzar periòdicament per tal de sincronitzar criteris. Aquesta sincronització continuada o periòdica s'ha evidenciat com essencial per mantenir la coherència del sistema.
- I la tercera és l'automatització. Alguns elements del sistema de governança són automatitzables, com el mapatge de dades (Alserafi et al., 2020; Rajmohan et al., 2023). La inclusió d'algoritmes d'intel·ligència artificial, ja en procés al Consorci d'Educació de Barcelona, permetrà a pocs anys vista automatitzar una part important de les tasques descrites a la **GMAP** i a la **GDAD**.

Un segon punt feble a destacar és que la identificació de les entitats no deixa de ser convencional, i per tant està exposada no només a errors, sinó també a biaixos. Els tres tipus d'incidències més habituals en aquest sentit són:

- Les errades en les descripcions de les entitats, o les identificacions duplicades, per manca de comunicació o sincronització entre els professionals dels diferents equips. L'ús de visualitzacions de dades per part d'aquests professionals, i el control periòdic, ajuden a reduir el nombre d'errades.

- La distorsió del criteri d'identificació. Les diferents entitats del sistema evolucionen amb el temps. El manteniment d'unes regles inicials d'identificació, si no s'actualitzen i es va actualitzant periòdicament, pot provocar identificacions poc coherents. La presa de decisions operatives, que sovint s'afegeix a les regles inicials creant doctrina, també pot esdevenir un factor de distorsió si no es revisa periòdicament.
- Determinades concepcions personals dels equips poden introduir biaixos ideològics en la identificació, especialment evidents en el cas de les dades i els objectes. Les revisions periòdiques dels mapes en la seva globalitat, juntament amb la incorporació de visions externes, permetrà detectar i esmenar aquests biaixos.

Voldríem posar el focus en un punt feble en concret, atès que es tracta d'un dels punts menys treballats en les metodologies arxivístiques. Es tracta de la identificació de les dades elementals, i la seva caracterització i assignació a objectes, un element que ha evidenciat una complexitat no prevista inicialment. En disposar de pocs referents metodològics directes (Ball, 2011; Dibowski et al., 2020; Sawadogo & Darmont, 2021; Wieder & Nolte, 2022), aquesta identificació ha tingut un component experimental, i s'ha anat depurant mitjançant revisions iteratives a partir de l'anàlisi de les mètriques del mapa de dades que s'anava composant. Els tres factors de complexitat han estat:

- **La identificació de la dada.** La dada elemental és una abstracció a partir d'una referència obtinguda d'una evidència o document, o de les interfícies o model de dades d'un sistema d'informació. Una mateixa dada pot aparèixer doncs amb denominacions diferents, i en funció del context generar molts dubtes en la seva abstracció. El codi d'un centre educatiu, per exemple, és una dada fàcilment identificable, perquè sempre apareixerà amb una denominació intuïtiva ("Codi centre", "ID_centre", etc.) i amb uns valors constants. Però el nom del centre, per exemple, pot presentar més dubtes. En un formulari de sol·licitud d'increment de rati pot aparèixer un camp textual denominat "Nom", que pel contingut podrem determinar que es tracta del nom del centre sol·licitant. En un altre formulari de comunicació d'inici d'activitat empresarial podem trobar el camp "Raó social", que correspon novament al nom del centre educatiu que demana legalitzar la seva activitat. I en un formulari on un docent sol·licita un certificat de serveis prestats, podem trobar, al costat del camp "Temps treballat", la denominació "Centre", que novament correspon al nom del centre educatiu. Per tant, existeix el risc de recollir com a tres dades elementals diferents el que correspon, en la seva abstracció, a la mateixa dada elemental. S'ha intentat eliminar al màxim aquestes duplicitats, però hem de tenir present que successives revisions podrien millorar la identificació, per la qual cosa el nombre de dades elementals identificades serà lleugerament superior al que correspondria a una identificació més precisa.

- **L'atribució del tipus de dada.** El tipus de dada no és una atribució absoluta, i en funció del context pot variar. Així doncs, una data de canvi d'estat d'un expedient s'ha considerat una dada "Transaccional", però la data de finalització del contracte d'un treballador/a s'ha considerat una dada "Laboral". De la mateixa manera, la descripció de l'estat de salut d'un alumne es pot considerar una dada "Descriptiva", però també una dada "Sanitària". S'observa doncs un potencial de millora important en la definició de criteris de categorització de les dades, al qual probablement podrà contribuir l'ús regular per a l'explotació del mapa de dades per part de grups de recerca i altres col·lectius.
- **La identificació de l'objecte descrit per la dada.** Aquesta ha estat una de les constatacions més rellevants i no previstes en el disseny inicial del projecte: a partir de la identificació del que inicialment s'havia considerat únicament un atribut de l'entitat "Dada elemental", s'ha evidenciat que a partir d'aquesta identificació d'objectes s'obria la possibilitat de construir un mapa de la part del món real descrita pel mapa de dades. I que, posant en relació els diferents objectes identificats, es podia construir progressivament una ontologia descriptiva del substrat informacional del Consorci. Voluntàriament s'ha limitat aquesta possibilitat treballant únicament amb dos nivells taxonòmics: l'objecte i el superobjecte. Identificar l'objecte descrit per una dada quan els objectes són coses del món real no ha plantejat problemes (un nom de centre descriu un centre educatiu). El principal problema en aquesta identificació ha estat determinar els objectes corresponents a accions, aquells que en podríem dir transaccionals, especialment quan l'acció produeix quelcom que després té existència en el món real. Posem un exemple. La tramitació d'un conveni és un objecte del tipus acció (un objecte transaccional), que produeix com a resultat un objecte del tipus cosa: el conveni. La data de signatura del conveni és al mateix temps la data de finalització de l'acció (una dada elemental transaccional descriptiva de la tramitació d'un conveni) i la data d'inici de vigència del conveni (una dada elemental descriptiva del conveni). No s'ha pogut formular una regla inequívoca per a resoldre aquestes variants, per la qual cosa s'ha optat per solucions puntuals convencionals. Es considera doncs que s'obre un camp de recerca rellevant en aquest aspecte.

8.2.2 Punts forts

La metodologia utilitzada ha post de relleu alguns beneficis directes sobre la governança de la informació de l'organització.³¹³

El principal benefici prové del coneixement que proporciona sobre els diferents elements del sistema de governança, el qual accelera la planificació de canvis, i permet estendre'ls a diferents processos en una mateixa actuació o projecte, valorant l'impacte de qualsevol canvi en d'altres processos, sistemes d'informació i sèries documentals. És per tant un factor accelerador de la transformació digital.

Aquest benefici es tradueix, a nivell operatiu, amb guanys concrets en àrees específiques de treball:

- En l'àmbit de la gestió documental, la metodologia permet una major precisió en la identificació de sèries documentals i la fixació de criteris d'arxiu, reduint el nombre de duplicitats o solapaments. També permet homogeneïtzar els criteris de nominalització dels expedients i utilitzar els nominals com a punts d'accés transversals.
- En l'àmbit dels processos, facilita la prioritització d'actuacions de transformació, en funció de l'impacte previst sobre la resta de processos i de la criticitat de cada actuació.
- En l'àmbit dels sistemes d'informació, permet prioritzar amb precisió els desenvolupaments tecnològics a programar, les tecnologies a utilitzar, i les integracions entre solucions que caldrà activar, ajustant el cost i l'esforç dels desenvolupaments en funció de la criticitat de cada procés.
- Així mateix, el treball sobre un mapa de dades global permet modelatge de dades més racional en el desenvolupament de sistemes, i la identificació i evolució progressiva de les taules mestres l'organització.

Així mateix, la metodologia afavoreix la coordinació dels equips tècnics implicats en la reenginyeria de processos i el desenvolupament de solucions tecnològiques, i n'incrementa la productivitat. L'ús dels mapes amb visualització interactiva s'ha revelat com una solució efectiva que facilita la coordinació dels equips de treball i millora la presa de decisions estratègiques a partir de la visió global que proporcionen.

³¹³ Valorem únicament els beneficis de la metodologia, no els beneficis que la dadificació dels processos aporta a una organització de forma general.

8.2.3 Recomanacions

Per tal d'assolir aquests beneficis, hi ha alguns trets de la metodologia que és especialment important tenir en compte:

- És important que la governança tingui un enfocament integral. La vocació des de l'inici del projecte ha de ser incloure totes les entitats de tots els àmbits implicats, sense exclusió, per poder aprofitar els beneficis del plantejament global.
- Les identificacions d'entitats sempre tenen un component utilitari. En tots els casos (processos, sèries, sistemes i dades) la realitat es pot entendre com un tot que es desenvolupa de forma contínua, sense solució en aquesta continuïtat. Els límits es posen de forma convencional, i en obediència a interessos estratègics o operatius concrets, per la qual cosa una mateixa realitat organitzacional pot ser descrita amb models i estructures diferents. Cal tenir present doncs que la coherència de la lectura global està per sobre dels límits inicials d'un procés o un sistema en concret, i que aquests es poden modificar en favor d'aquesta coherència.
- És important desenvolupar el model de governança fins al nivell més baix d'agregació, el nivell de dada, atès que bona part dels beneficis que pot proporcionar el model s'obtenen a partir de mesurar el paper i la presència de les dades en els processos, les sèries i els sistemes d'informació. Sense arribar al nivell de dada, el sistema perd bona part del seu valor com a eina de governança.
- La descomposició de la dada elemental en les seves successives ocurrencies és especialment rellevant per traçar el llinatge de la dada, no només entre sistemes d'informació, sinó també entre evidències. El control del llinatge de la dada és un element clau per poder determinar la fiabilitat i autenticitat de la dada.
- És impossible desplegar d'un sol cop el sistema de governança, ni tampoc arribar a un estadi que es pugui considerar definitiu. Les organitzacions i els seus components informacionals evolucionen i es modifiquen contínuament, i per tant el sistema de governança s'ha d'anar desplegant i actualitzant al mateix ritme. L'organicitat del disseny del sistema³¹⁴ és clau doncs per garantir-ne la seva eficàcia i actualitat.
- I finalment la capacitat d'establir mesures i indicadors quantitius proporciona un element de control determinant, al mateix temps que permet definir alertes automatitzades per al seguiment de l'evolució dels mapes. L'ús d'eines de visualització de dades és un element de gran ajuda en aquesta monitorització.

³¹⁴ Entesa com el creixement orgànic o natural dels mapes.

8.2.4 Futures línies de recerca

El potencial demostrat per la metodologia proposada ens permet apuntar futures línies de recerca en l'àmbit de la governança de la informació i la gestió documental, que considerem que són les següents:

- Validar la utilitat de la metodologia per a la millora de la qualitat d'un sistema de gestió documental, especialment en allò que fa referència a la identificació i classificació de sèries, a partir d'operacions d'homogeneïtzació de sèries documentals i d'optimització del substrat documental dels processos.
- Validar la utilitat de la metodologia per a la priorització i acceleració de les accions de transformació digital o millora de processos i per assolir un nivell alt de madificació en l'organització.
- Validar la utilitat de la metodologia com una eina per a la planificació del mapa de sistemes d'una organització de forma coordinada amb la visió de la gestió documental, i per a la priorització del desenvolupament de nous sistemes d'informació o la transformació, evolució o integració de sistemes ja existents.

8.3 Epíleg: resum de les conclusions generals

8.3.1 Objectiu 1: característiques de la sèrie documental dadificada

La sèrie documental dadificada es caracteritza per disposar d'un nombre elevat de dades disponibles en format estructurat, un 64% de mitjana segons l'anàlisi realitzada en el capítol 7, mentre que gairebé un 50% de les evidències corresponen de mitjana a evidències basades en dades.

La dadificació es manifesta més efectiva en els processos on només participen agents interns, atès que permet un desplegament ràpid i integral utilitzant eines de desenvolupament *low code*. La participació d'agents externs a l'organització incrementa el nombre d'evidències-objecte en la sèrie documental i redueix les possibilitats d'una dadificació completa, a causa de la complexitat de donar accés als agents externs.

Finalment, s'ha observat una certa regularitat en la proporció d'objectes de tipus "cosa" i "acció" que apareixen en les sèries documentals. Aquesta proporció és aproximadament de 3 a 1, amb una forquilla que va de 2 a 4 objectes "cosa" per cada objecte "acció".

8.3.2 Objectiu 2: característiques de l'ecosistema tecnològic de la sèrie documental dadificada

En les organitzacions similars a la que ha estat objecte d'estudi, l'arquitectura de sistemes d'informació utilitzats per a la gestió segueix en general un model tradicional, basat en aplicacions desenvolupades sobre bases de dades relacionals, i en alguns casos amb solucions MDM per a la consolidació de les dades mestres. En un model tradicional, la informació que no queda coberta per aplicacions es gestiona amb ofimàtica, i queda per tant fora de les oportunitats de dadificació.

El model alternatiu que facilita accelerar la dadificació de qualsevol tipus de procés de l'organització és el model *data lake* o *data fabric*, el qual ha estat utilitzat de forma preferent en el cas d'aplicació del Consorci per a dadificar processos. Aquest model permet combinar fonts d'informació de diferents entorns tecnològics, incloent l'ofimàtica, i fa possible una governança integrada com la que s'assoleix amb una aplicació

informàtica tradicional. Aquest model es combina amb l'ús d'entorns de col·laboració, desenvolupament lleuger d'aplicacions (*low code*) i automatitzacions externes a les aplicacions (RPA).

8.3.3 Objectiu 3: equivalència entre la sèrie documental i el sistema d'informació

No existeix una equivalència natural entre sèrie documental i sistema d'informació. De la mateixa manera, un *dataset* o conjunt de dades no es pot considerar per sí sol una forma documental, i tindrà una consideració o una altra en funció del context que el doti d'una determinada significació. I aquest context dependrà de decisions preses en la fase de disseny, tant pels equips de desenvolupament d'aplicacions com pels equips de gestió documental, i documentades posteriorment.

Un *dataset* es pot ubicar en diferents punts de l'estructura d'un sistema d'informació. En funció d'aquesta ubicació, i de la capacitat d'adquirir context, s'ha proposat unes potencials equivalències del *dataset* amb entitats documentals, que es detallen en la Taula 8-2, i que es poden resumir en l'equivalència entre el conjunt de dades, la matriu de variables i la sèrie primària de tot procés.

8.3.4 Objectiu 4: les arquitectures distribuïdes com a entorns de gestió (transaccionals)

El cas d'aplicació ha confirmat quantitativament que es poden gestionar processos íntegrament basats en dades només amb tecnologia *low code*, disponible en entorns de col·laboració basats en bases de dades orientades a objectes, emmagatzemant les dades en un entorn distribuït tipus *data fabric*, i completant la gestió amb visualitzacions interactives de dades. Per aquest motiu, es considera que els sistemes distribuïts són vàlids com entorn de gestió, i per tant poden ser considerats també transaccionals.

8.3.5 Objectiu 5: equivalència entre tipus de sèries documentals i tipus de sistemes d'informació

No existeix una relació evident ni natural entre tipus específics de sèrie documental i tipus de sistemes d'informació. Com s'ha analitzat en l'objectiu 3, tant les sèries documentals, com especialment els sistemes d'informació, són un producte de decisions preses en la fase de disseny, que poden estar inspirades per criteris diferents al funcional (criteris d'oportunitat, de dificultat o risc, de cost, de preferències metodològiques, etc.). Però per aquest mateix motiu, convenim que, si els equips de disseny així ho preveuen, és possible fer coincidir conceptualment la sèrie documental i el sistema d'informació.

A part d'això, no s'ha pogut identificar en les sèries dadificades les tipologies tradicionals de sèrie documental (transaccional, recapitulativa i factícia). Totes les sèries dadificades tenen caràcter transaccional, i estructuralment es poden considerar recapitulatives (en forma registral), per la qual cosa la diferència desapareix. Però es pot reconsiderar a partir dels objectes descrits per la sèrie, possibilitat que es desenvolupa en l'objectiu 7.

8.3.6 Objectiu 6: les visualitzacions de dades com a recurs per a la gestió

La introducció de les visualitzacions interactives de dades com un element més d'una sèrie documental, fent la doble funció d'evidència o document de la sèrie i d'eina de gestió, incrementa de forma directa el nombre de dades de la sèrie documental, el nombre d'objectes dels quals gestiona informació, i les possibilitats de la sèrie d'intercanviar dades amb altres sèries documentals. Es considera doncs una pràctica recomanable incorporar les visualitzacions de dades com a forma documental, i una línia de recerca a explorar.

8.3.7 Objectiu 7: de la classificació funcional a la classificació ontològica

Les entitats funcionals (els processos o activitats i les sèries que se'n deriven), objecte tradicional de la classificació funcional, contenen informació sobre dos tipus d'objectes: "accions" i "coses". Aquesta constatació duplica el criteri de classificació que es pot

aplicar a les dades d'aquestes entitats, i qüestiona la validesa del model funcional com a sistema únic de classificació arxivística.

En les sèries i evidències, ambdós tipus d'objectes descrits es combinen formant sentències. La necessitat de capturar, a més de les dades, aquestes relacions causals entre les dades, obre la porta a utilitzar models ontològics per a classificar les sèries documentals en els entorns dadificats.

En aquest sentit, es proposa utilitzar la pròpia estructura dels entorns distribuïts (*data lake* o *data fabric*) per combinar el sistema de classificació funcional tradicional amb un sistema de classificació ontològic, aplicant el següent patró d'actuació:

- A. Considerar l'existència d'un tipus de sèrie documental sempre transaccional en la capa d'ingesta, que serà el resultat de la generació de vistes a partir dels sistemes d'informació de gestió.
- B. Considerar un segon tipus de sèrie documental optimitzada, representada pels dominis del *business glossary*, que podrà ser de dos tipus: relativa a "accions", on es mantindrà, agregat, el criteri funcional, o relativa a "coses" del món real.
- C. Considerar les vistes generades amb finalitats d'explotació com elements temporals que, en cas d'adquirir valor evidencial, s'incorporaran a la sèrie transaccional corresponent en la capa d'ingesta.

En conclusió, i com s'ha pogut demostrar en l'apartat 7.8 del capítol 7, es pot afirmar que:

- La classificació funcional és aplicable a un entorn dadificat (i per tant permet classificar sistemes d'informació i conjunts de dades generats per aquests sistemes).
- La classificació funcional i els models ontològics són compatibles en un mateix entorn de gestió dadificada, ubicats en la capa de manteniment d'un sistema d'emmagatzemament distribuït, cosa la qual permet optimitzar l'aplicació de les polítiques d'arxiu.

Capítol 9. Bibliografia³¹⁵

Abedjan, Z., Golab, L., & Naumann, F. (2016). *Data profiling*. IEEE 32nd International Conference on Data Engineering (ICDE), Helsinki, Finland.
<https://doi.org/10.1109/ICDE.2016.7498363>

Acland, G. (2006). The Australian Recordkeeping Metadata Schema—Version 1.0: Note from the Research Team. *Archivaria*, 49.

Aguayo, I. M. (2021). *Ontology of Basque Government's Records Management Model (first version) / Ontología del Modelo de Gestión Documental del Gobierno Vasco (primera version)*. <https://doi.org/10.5281/ZENODO.5504368>

Akerreta, L., & Moyano, J. (2021). Contar historias con los datos: Aragón Open Data Focus, una experiencia innovadora de reutilización de los datos del sector público. *Scire: representación y organización del conocimiento*, 27(1).
<https://doi.org/10.54886/scire.v27i1.4777>

Alalwan, J., & Thomas, M. (2011). Designing ERM ontology to evaluate records management system. *Proceedings of the Annual Hawaii International Conference on System Sciences*. <https://doi.org/10.1109/HICSS.2011.149>

Alalwan, J., & Thomas, M. (2012). An Ontology-based Approach to Assessing Records Management Systems. *E-Service Journal*, 8(3).

Alamillo, I. (2016). La regulación de la identificación, la firma y sello electrónicos en la Ley 39/2015. *Consultor de los ayuntamientos y de los juzgados: Revista técnica especializada en administración local y justicia municipal*, 24.

Alamillo, I. (2019). El RDL 14/2019: Una extraordinariamente urgente e innecesaria reforma del régimen de identificación y firma electrónica en la LPAC. *Consultor de los ayuntamientos y de los juzgados: Revista técnica especializada en administración local y justicia municipal*, 12.

³¹⁵ Per a la bibliografia general s'ha utilitzat el format de citació bibliogràfica *American Psychological Association (APA) 7th Edition*.

Les referències a normes de descripció arxivística i models de descripció equivalents no s'inclouen en aquesta bibliografia, atès que se citen en l'apartat 4.5 (Models de referència i abreviatures per citació) del capítol 4. Les abreviatures fixades en l'apartat 4.5 s'utilitzen per a fer referència a aquestes normes de descripció en tot el text de la tesi.

Les referències a pàgines Web se citen en nota a peu de pàgina.

Alamillo, I. (2021). La identificación y firma electrónica en el Real Decreto 203/2021 de estos polvos, estos lodos.... *Consultor de los ayuntamientos y de los juzgados: Revista técnica especializada en administración local y justicia municipal*, 5.

Alamillo, I., Balloriani, M., Bernal, M. Á., & Gallone, G. (2021). The future of public administration through the use of blockchain technology. *European Review of Digital Administration & Law*, 2(2).

Alberch, R. (2016, juliol 13). La nueva ISO 15489:2016. Éramos pocos y parió la abuela. *Factor GDA*. <https://esaged.wordpress.com/2016/07/13/la-nueva-iso-154892016-eramos-pocos-y-pario-la-abuela/>

Alberch, R., Boadas, J., Capell, E., & Corominas, M. (2009). *Manual d'arxivística i gestió documental*. Associació d'Arxivers de Catalunya.

Aldama, O., Delgado, M., & Díaz-Canel, M. (2022). Metodología de los tableros y cuadro de mando integral en la gestió de gobierno orientada a la innovación. *Revista Cubana de Administración Pública y Empresarial*, 6(3).
<https://doi.org/10.5281/ZENODO.7062669>

Alonso, J. A., & Lloveras, M. R. (2007). *Aproximación a la gestión de documentos desde el enfoque basado en procesos*. 10as Jornadas Españolas de Documentación, Santiago de Compostela.

Alserafi, A., Abelló, A., Romero, O., & Calders, T. (2020). Keeping the Data Lake in Form: Proximity Mining for Pre-Filtering Schema Matching. *ACM Transactions on Information Systems*, 38(3). <https://doi.org/10.1145/3388870>

Anderson, B., Eaton, F., & Schwartz, S. W. (2015). Archival appraisal and the digital record: Applying past tradition for future practice. *New Review of Information Networking*, 20(1). <https://doi.org/10.1080/13614576.2015.1114823>

Aparecida, L., Pelogia, I., & Inês Dallavalle, S. (2012). Process management tasks and barriers: Functional to processes approach. *Business Process Management Journal*, 18(5). <https://doi.org/10.1108/14637151211270144>

Armbrust, M., Ghodsi, A., Xin, R., & Zaharia, M. A. (2021). *Lakehouse: A New Generation of Open Platforms that Unify Data Warehousing and Advanced Analytics*. Conference on Innovative Data Systems Research.

Attard, J., Orlandi, F., Scerri, S., & Auer, S. (2015). A systematic review of open government data initiatives. *Government Information Quarterly*, 32(4).
<https://doi.org/10.1016/j.giq.2015.07.006>

Bagozi, A., Bianchini, D., De Antonellis, V., Garda, M., & Melchiori, M. (2019). Personalised Exploration Graphs on Semantic Data Lakes. A H. Panetto, C. Debruyne, M. Hepp, D. Lewis, C. A. Ardagna, & R. Meersman (Ed.), *On the Move to Meaningful*

Internet Systems: OTM 2019 Conferences (Vol. 11877). Springer International Publishing. https://doi.org/10.1007/978-3-030-33246-4_2

Bailey, S. (2009). Forget electronic records management, it's automated records management that we desperately need. *Records Management Journal*, 19(2).

Bak, G. (2010). *Impacts of Web 2.0 on Information Models: Life Cycle and Continuums*. Library and Archives Canada.

<https://www.collectionscanada.gc.ca/obj/012018/f2/012018-3403-e.pdf>

Bak, G. (2012). Continuous classification: Capturing dynamic relationships among information resources. *Archival Science*, 12(3). <https://doi.org/10.1007/s10502-012-9171-8>

Bak, G. (2016a). Not Meta Just Data: Redefining Content and Metadata in Archival Theory and Practice. *Journal of Archival Organization*, 13(1-2).

<https://doi.org/10.1080/15332748.2017.1413974>

Bak, G. (2016b). Trusted by whom? TDRs, standards culture and the nature of trust. *Archival Science*, 16(4). <https://doi.org/10.1007/s10502-015-9257-1>

Ball, C. (2011, setembre 23). A Bit About Data Mapping. *Ball in your Court: Musings on e-Discovery & Computer forensics*.

<https://ballinyourcourt.wordpress.com/2011/09/23/a-bit-about-data-mapping/>

Bantin, P. C. (1998). Strategies for managing electronic records: A new archival paradigm? An affirmation of our archival traditions? *Archival Issues*, 23(1).

Barbadillo, J. (2010). Clasificaciones y relaciones funcionales de los documentos de archivo. *Tabula: revista de archivos de Castilla y León*, 13.

Barros, T. H. B., & Sousa, R. T. B. D. (2019). Archival Science and Knowledge Organization: Mapping Methodological Relationships. *Knowledge Organization*, 46(7).

<https://doi.org/10.5771/0943-7444-2019-7-493>

Beach, J., & Oates, J. (2014). Maintaining best practice in record-keeping and documentation. *Nursing Standard*, 28(36).

<https://doi.org/10.7748/ns2014.05.28.36.45.e8835>

Bearman, D. (1991). *Archival management of electronic records* (13; Archives and Museum Informatics Technical Report).

Bearman, D. (1994). Electronic evidence: Strategies for managing records in contemporary organizations. *Archives & Museums Informatics*, 19.

- Bearman, D. (2006). Moments of risk: Identifying threats to electronic records. *Archivaria*, 62.
- Becker, C. (2018). Metaphors we work by: Reframing digital objects, significant properties, and the design of digital preservation systems. *Archivaria*, 85.
- Beheshti, A., Benatallah, B., Nouri, R., & Tabebordbar, A. (2018). CoreKG: A knowledge lake service. *Proceedings of the VLDB Endowment*, 11(12). <https://doi.org/10.14778/3229863.3236230>
- Beheshti, A., Benatallah, B., Sheng, Q. Z., & Schiliro, F. (2020). Intelligent Knowledge Lakes: The Age of Artificial Intelligence and Big Data. A L. H. U, J. Yang, Y. Cai, K. Karlapalem, A. Liu, & X. Huang (Ed.), *Web Information Systems Engineering* (Vol. 1155). Springer International Publishing. https://doi.org/10.1007/978-981-15-3281-8_3
- Beheshti, A., Yakhchi, S., Mousaeirad, S., Ghafari, S. M., Goluguri, S. R., & Edrisi, M. A. (2020). Towards Cognitive Recommender Systems. *Algorithms*, 13(8). <https://doi.org/10.3390/a13080176>
- Bell, M. (2020). From tree to network: Reordering an archival catalogue. *Records Management Journal*, 30(3). <https://doi.org/10.1108/RMJ-09-2019-0051>
- Bellatreche, L., Goyal, V., Fujita, H., Mondal, A., & Reddy, P. K. (2020). *Big Data Analytics*. 8th International Conference BDA, Sonapat.
- Benfeldt, O., Persson, J. S., & Madsen, S. (2020). Data Governance as a Collective Action Problem. *Information Systems Frontiers*, 22(2). <https://doi.org/10.1007/s10796-019-09923-z>
- Borglund, E., & Engvall, T. (2014). Open data?: Data, information, document or record? *Records Management Journal*, 24(2). <https://doi.org/10.1108/RMJ-01-2014-0012>
- Boss, K., & Broussard, M. (2017). Challenges of archiving and preserving born-digital news applications. *IFLA Journal*, 43(2). <https://doi.org/10.1177/0340035216686355>
- Bouth, F., & Moreira, C. (2022). Classificação arquivística: Uma perspectiva da metodologia funcional vinculada ao tipo documental. *Scire: representación y organización del conocimiento*, 28(2). <https://doi.org/10.54886/scire.v28i2.4802>
- Bowker, G. C., & Star, S. L. (1999). *Sorting things out: Classification and its consequences*. MIT Press.
- Bowker, L., & Villamizar, C. (2017). Embedding a records manager as a strategy for helping to positively influence an organization's records management culture. *Records Management Journal*, 27(1). <https://doi.org/10.1108/RMJ-02-2016-0005>

Boyd, D., & Crawford, K. (2012). Critical questions for Big Data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 15(5). <https://doi.org/10.1080/1369118X.2012.678878>

Brandes, U., Eiglsperger, M., Herman, I., Himsolt, M., & Marshall, M. S. (2002). GraphML Progress Report Structural Layer Proposal. En P. Mutzel, M. Jünger, & S. Leipert (Ed.), *Graph Drawing* (Vol. 2265). Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-45848-4_59

Brooks, J. (2019). Perspectives on the relationship between records management and information governance. *Records Management Journal*, 29(1/2). <https://doi.org/10.1108/RMJ-09-2018-0032>

Buckland, M. (1991). Information as Thing. *Journal of the American Society for Information Science*, 42(5).

Buckland, M. (1994). On the Nature of Records Management Theory. *American Archivist*, 57(2).

Bullivant, R. (2018, maig 29). Data catalogs might be the new black, but metadata discovery to provision them can be tricky. *Data catalog, Data Governance, Enterprise Metadata Management*. <https://www.silwoodtechnology.com/blog/data-catalogues-might-be-the-new-black-but-metadata-discovery-to-provision-them-can-be-tricky/>

Burghof, H. (2000). *Managing records in multi-system organizations: An In-Place Approach to Information Governance* (52). RecordLion.

Burk, C. F. (Cornelius F., 1933-, & Horton, F. W. (1988). *InfoMap: A complete guide to discovering corporate information resources*. Prentice-Hall.

Bustelo, C. (2011). Norma ISO 30300. Management system for records. *Anuario ThinkEPI*, 5.

Bustelo, C. (2012). La normalización internacional en información y documentación: ¿una historia de éxitos? El caso de la normalización ISO en gestión de documentos. *Métodos de informacion*, 3(4). <https://doi.org/10.5557/IIMEI2-N2-039046>

Butler, D. (2011). *Master Data Management* (An Oracle White Paper). Oracle Corporation.

Campos, P. (2023). Conjunts de dades, gestió documental i arxius: Proposta de tractament. *Lligall; Revista catalana d'arxivística*, 46. <https://raco.cat/index.php/liligall/article/view/416822>

Caplan, P. (2021). *Understanding PREMIS* (3a versió revisada). Library of Congress Network Development and MARC Standards Office.

Caplinger, C., & Borgschulte, A. (2015). *Information-Governance-Assessment-Guide*. RecordLion.

Casellas, L.-E. (2014). The mapping, selecting and opening of data: The records management contribution to the Open Data project in Girona City Council. *Records Management Journal*, 24(2). <https://doi.org/10.1108/RMJ-01-2014-0008>

Casellas, L.-E. (2019). L'avaluació arxivística en el nou context de les organitzacions. En *La destrucció d'informació pública: Una mirada multidisciplinària sobre l'eliminació ordenada de la documentació* (1a ed.). Oficina Antifrau de Catalunya; Associació d'Arxivers-Gestors de Documents de Catalunya.

Chandler, K. (2022). Reevaluating the mission of control in records management: Tools for a staff-centered approach. *Records Management Journal*, 32(3). <https://doi.org/10.1108/RMJ-05-2022-0012>

Chen, Z. (2022). Observations and Expectations on Recent Developments of Data Lakes. *Procedia Computer Science*, 214. <https://doi.org/10.1016/j.procs.2022.11.192>

Chessa, A., Fenu, G., Motta, E., Osborne, F., Salatino, A., & Secchi, L. (2022). *Enriching Data Lakes with Knowledge Graphs*. Text2KG: International Workshop on Knowledge Graph Generation from Text, Hersonissos, Crete.

Chinosi, M., & Trombetta, A. (2012). BPMN: An introduction to the standard. *Computer Standards & Interfaces*, 34(1). <https://doi.org/10.1016/j.csi.2011.06.002>

Chorley, K. M. (2017). The challenges presented to records management by open government data in the public sector in England: A case study. *Records Management Journal*, 27(2). <https://doi.org/10.1108/RMJ-09-2016-0034>

Cisco, S. (2008a). Big buckets for simplifying records retention schedules. *Information Management Journal*, 42(5). Gale Academic OneFile.

Cisco, S. (2008b). How to win the compliance battle using «big buckets». *Information Management Journal*, 42(4). Gale Academic OneFile.

Clavaud, F., & Wildi, T. (2021). *ICA Records in Contexts-Ontology (RiC-O): A Semantic Framework for Describing Archival Resources*. Proceedings of Linked Archives International Workshop co-located with 25th International Conference on Theory and Practice of Digital Libraries (TPDL 2021). <https://ceur-ws.org/Vol-3019/>

Cloonan, M. V., Gilliland-Swetland, A., Marciano, R., Thibodeau, K., & Trace, C. (2005). InterPARES 2: Experiential, interactive, and dynamic records. *Proceedings of the American Society for Information Science and Technology*, 39(1). <https://doi.org/10.1002/meet.1450390161>

- Codd, E. F. (1970). A relational model of data for large shared data banks. *Communications of the ACM*, 13(6). <https://doi.org/10.1145/362384.362685>
- Codd, E. F. (1971). *Normalized data base structure: A brief tutorial*. Proceedings of the ACM SIGFIDET (now SIGMOD) Workshop on Data Description, Access and Control - SIGFIDET '71, San Diego, California. <https://doi.org/10.1145/1734714.1734716>
- Committee on Electronic Records. (1997). *Guide for Managing Electronic Records from an Archival Perspective* (Núm. 8; ICA Study). International Council on Archives.
- Convery, N. (2014). From reactive to proactive appraisal. *Archives and Manuscripts*, 42(2).
- Convery, N. (2016, july 27). There is a new version of the international standard for records management (ISO 15489-1:2016). *JISC Infrastructure*. <https://infrastructure.jiscinvolve.org/wp/2016/06/27/there-is-a-new-version-of-the-international-standard-for-records-management-iso-15489-12016/>
- Cook, T. (2013). Evidence, memory, identity, and community: Four shifting archival paradigms. *Archival Science*, 13(2-3). <https://doi.org/10.1007/s10502-012-9180-7>
- Costa, L. F., & Ramalho, J. C. (2021). *Towards Entity Linking, NER in Archival Finding Aids*. Proceedings of Linked Archives International Workshop co-located with 25th International Conference on Theory and Practice of Digital Libraries (TPDL 2021). <https://ceur-ws.org/Vol-3019/>
- Cottin, M., & Dessolin, S. (2012). La famille des normes ISO sur le records management. *La Gazette des archives*, 228(4). <https://doi.org/10.3406/gazar.2012.4989>
- Craig, B. (2004). *Archival appraisal: Theory and practice*. K.G. Saur.
- Cumming, K., & Picot, A. (2014). Reinventing appraisal. *Archives and Manuscripts*, 42(2). <https://doi.org/10.1080/01576895.2014.926824>
- Cunningham, A., Millar, L. A., & Reed, B. (2013). Peter J. Scott and the Australian 'series' system: Its origins, features, rationale, impact and continuing relevance. *Comma*, 1. <https://doi.org/10.3828/comma.2013.1.13>
- Currie, M. E., & Paris, B. S. (2018). Back-ups for the future: Archival practices for data activism. *Archives and Manuscripts*, 46(2). <https://doi.org/10.1080/01576895.2018.1468273>

Dahlberg, I. (1993). Knowledge organization: Its scope and possibilities. *Knowledge Organization*, 20(4). <https://doi.org/10.5771/0943-7444-1993-4-211>

Davenport, T. H., & Prusak, L. (1998). *Working Knowledge: How Organizations Manage what They Know*. Harvard Business School Press.

De Man, H. (2009). Case management: A review of modeling approaches. *BPTrends*, January, 2009.

Dehghani, Z. (2019). How to Move Beyond a Monolithic Data Lake to a Distributed Data Mesh. *Martinfowler.Com*. <https://martinfowler.com/articles/data-monolith-to-mesh.html>

Delgado, A. (2007). La indeterminación de la traducción archivística. *El Profesional de la Información*, 16(1).

Delgado, A. (2010). Sistemas de clasificación en múltiples dimensiones: La experiencia del Archivo Municipal de Cartagena. *Tabula: revista de archivos de Castilla y León*, 13.

Delgado, T. (2022). Arquitectura de referencia de ecosistemas de datos basada en data mesh & data fabric. *Revista Cubana de Administración Pública y Empresarial*, 6(3).

Diamantini, C., Giudice, P. L., Musarella, L., Potena, D., Storti, E., & Ursino, D. (2018). A New Metadata Model to Uniformly Handle Heterogeneous Data Lake Sources. A A. Benczúr, B. Thalheim, T. Horváth, S. Chiusano, T. Cerquitelli, C. Sidló, & P. Z. Revesz (Ed.), *New Trends in Databases and Information Systems* (Vol. 909). Springer International Publishing. https://doi.org/10.1007/978-3-030-00063-9_17

Diamantini, C., Lo Giudice, P., Potena, D., Storti, E., & Ursino, D. (2021). An Approach to Extracting Topic-guided Views from the Sources of a Data Lake. *Information Systems Frontiers*, 23(1). <https://doi.org/10.1007/s10796-020-10010-x>

Dibowski, H., Schmid, S., Svetashova, Y., Henson, C., & Tran, T. (2020). *Using Semantic Technologies to Manage a Data Lake: Data Catalog, Provenance and Access Control*. SSWS@ISWC.

Diedel, N. H. & M. (2020). Managing Unstructured Content for Transactional Systems. A *Gimnal*. <https://www.gimnal.com/videos/recording-webinar-managing-unstructured-content-for-transactional-systems?submissionGuid=07305784-f698-48fa-aeb9-fb770ef78b73>

Dixon, J. (2010). Pentaho, Hadoop, and Data Lakes. *James Dixon's Blog*. <https://jamesdixon.wordpress.com/2010/10/14/pentaho-hadoop-and-data-lakes/>

Duff, W. M., & Haskell, J. (2015). New Uses for Old Records: A Rhizomatic Approach to Archival Access. *American Archivist*, 78(1). <https://doi.org/10.17723/0360-9081.78.1.38>

Duranti, L. (1997). The Archival Bond. *Archives and Museum Informatics*, 11(3-4). <https://doi.org/10.1023/A:1009025127463>

Duranti, L. (1999). D'aquí a l'eternitat: Conceptes i principis de la gestió de documents electrònics. *Lligall; Revista catalana d'arxivística*, 14. <https://raco.cat/index.php/liligall/article/view/339483>

Duranti, L. (2016). *Comments on "Records in Context"*. InterPARES Trust. http://interparestrust.org/assets/public/dissemination/interparestrust_comments_on_ri_c_fi nal2.pdf

Duranti, L. (2022). Why a world gone digital needs archival theory more than ever before? *Archeion*, 123. <https://doi.org/10.4467/26581264ARC.22.015.16668>

Duranti, L., Eastwood, T., & MacNeil, H. (2002). *Preservation of the integrity of electronic records*. Springer International Publishing.

Duranti, L., & MacNeil, H. (1996). The protection of the integrity of electronic records: An overview of the UBC-MAS research project. *Archivaria*, 42.

Earley, S., & Henderson, D. (Ed.). (2017). *DAMA-DMBOK: Data management body of knowledge* (2nd edition). Data Management Association.

Efthymiou, V., Papadakis, G., Papastefanatos, G., Stefanidis, K., & Palpanas, T. (2017). Parallel meta-blocking for scaling entity resolution over big heterogeneous data. *Information Systems*, 65. <https://doi.org/10.1016/j.is.2016.12.001>

Ehrlinger, L., Schrott, J., Melichar, M., Kirchmayr, N., & Wöß, W. (2021). *Data Catalogs: A Systematic Literature Review and Guidelines to Implementation*. Database and Expert Systems Applications - DEXA Workshops.

Escobar, L. F. (2009). Tablas de retención documental enfocada a procesos (TRDP): Una propuesta emergente en tiempos de cambio. *Revista Códice*, 5(2).

Expanding the Use of a Role-Based Approach (Capstone) for Electronic Messages (Bulletin 2023-02). (2023). National Archives and Records Administration (NARA). <https://www.archives.gov/records-mgmt/bulletins/2023/2023-02>

Feilmayr, C., & Wöß, W. (2016). An analysis of ontologies and their success factors for application to business. *Data & Knowledge Engineering*, 101.

- Feldman, P. (2003). Applying the Australian Series system to the management of current records. *Records Management Journal*, 13(2).
<https://doi.org/10.1108/09565690310485270>
- Ferreira, M. (2006). *Introdução à preservação digital. Conceitos, estratégias e actuais consensos. Minho Portugal: Edição Eletrônica, Escola de Engenharia da Universidade do Minho.*
- Findlay, C. (2017). Participatory cultures, trust technologies and decentralisation: Innovation opportunities for recordkeeping. *Archives and Manuscripts*, 45(3).
<https://doi.org/10.1080/01576895.2017.1366864>
- Findlay, C. (2018). Crunch time: The revised ISO 15489 and the future of recordkeeping. *Archives and Manuscripts*, 46(2).
<https://doi.org/10.1080/01576895.2018.1451755>
- Findlay, C., & Cumming, K. (2021, abril 11). Recordkeeping theory, models & strategies and today's workplace. *Recordkeeping Roundtable*.
<https://rkroundtable.org/2021/04/11/recordkeeping-theory-models-strategies-and-todays-workplace/>
- Findlay, C., Cumming, K., Picot, A., & Reed, B. (2014). Reinventing Archival Methods. *Archives and Manuscripts*, 42(2).
- Fonseca, L., Amaral, A., & Oliveira, J. (2021). Quality 4.0: The EFQM 2020 Model and Industry 4.0 Relationships and Implications. *Sustainability*, 13(6).
<https://doi.org/10.3390/su13063107>
- Foscarini, F. (2006). Records classification and functions: An archival perspective. *Knowledge Organization*, 33(4).
- Foscarini, F. (2012). Understanding functions: An organizational culture perspective. *Records Management Journal*, 22(1).
- Foscarini, F., Krucker, M., & Golick, D. (2022). Meeting technologies and recordkeeping: A preliminary study. *Records Management Journal*, 32(1).
<https://doi.org/10.1108/RMJ-07-2021-0028>
- Francart, T., Clavaud, F., & Charbonnier, P. (2021). *RiC-O Converter: A Software to Convert EAC-CPF and EAD 2002 XML Files to RDF Datasets Conforming to Records in Contexts Ontology*. Proceedings of Linked Archives International Workshop co-located with 25th International Conference on Theory and Practice of Digital Libraries (TPDL 2021). <https://ceur-ws.org/Vol-3019/>
- Franco, B., & Pérez, R. (2015). *Archives Model of the Transparency and Access to Public Information Network (RTA)* (Vol. 30). EUROsocial Programme.

- Franks, P. C. (2013). *Records and information management*. Facet.
- Franks, P. C. (2020). Implications of blockchain distributed ledger technology for records management and information governance programs. *Records Management Journal*, 30(3). <https://doi.org/10.1108/RMJ-08-2019-0047>
- Fraser, K. C. (1999). Diplomatics: New Uses for an Old Science. *Reference Reviews*, 13(6).
- Frendo, R. (2007). Disembodied information: Metadata, file plans, and the intellectual organisation of records. *Records Management Journal*, 17(3). <https://doi.org/10.1108/09565690710833062>
- Frontoni, E., Paolanti, M., Lauriault, T. P., Stiber, M., Duranti, L., & Muhammad, A.-M. (2022). Trusted Data Forever: Is AI the Answer? *ArXiv*. <http://arxiv.org/abs/2203.03712>
- Galarreta, P. R. de, Cabrera, M. D., & Rodríguez, J. M. (2012). La teoría desde la práctica. Nuevas reflexiones sobre el concepto de serie documental. *Arch-e: Revista Andaluza de Archivos*, 5.
- Galloway, P. (2008). *Big Buckets or Big Ideas? Classification vs. Innovation on the Enterprise 2.0 Desktop*. ARMA International Educational Foundation; Mile High Denver Chapter of ARMA International.
- Gandomi, A., & Haider, M. (2014). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2). <https://doi.org/10.1016/j.ijinfomgt.2014.10.007>
- García Ruipérez, M. (2007). Series y tipos documentales: Modelos de análisis. *Legajos: cuadernos de investigación archivística y gestión documental*, 10.
- García-Morales, E. (2014). El mapa documental y la gestión documental orientada a procesos. *Anuario ThinkEPI*, 8(1).
- Garmendia, J., & Retter, A. (2021). *Developing a Pan-Archival Linked Data Catalogue*. Proceedings of Linked Archives International Workshop co-located with 25th International Conference on Theory and Practice of Digital Libraries (TPDL 2021). <https://ceur-ws.org/Vol-3019/>
- Ghavami, P. (2020). *Big data management: Data governance principles for big data analytics* (1a ed.). De Gruyter.
- Ghiran, A.-M., & Buchmann, R. A. (2019). The Model-Driven Enterprise Data Fabric: A Proposal Based on Conceptual Modelling and Knowledge Graphs. *A Knowledge Science, Engineering and Management*. Springer International Publishing.

Giebler, C., Gröger, C., Hoos, E., Eichler, R., Schwarz, H., & Mitschang, B. (2021). *The Data Lake Architecture Framework: A Foundation for Building a Comprehensive Data Lake Architecture*. Proceedings der 19. Fachtagung für Datenbanksysteme für Business, Technologie und Web (BTW), Stuttgart.

Gitelman, L. (Ed.). (2013). «*Raw Data*» *Is an Oxymoron*. MIT Press.
<https://doi.org/10.7551/mitpress/9302.001.0001>

Guardado, C. (2013). *A classificação da informação arquivística da administração local nos países ibéricos: Uma análise comparada*. Jornadas Ibéricas de Arquivos Municipais, Lisboa.

Gueguen, G., da Fonseca, V., Pitti, D., & Grimoüard, C. (2013). Toward an International Conceptual Model for Archival Description: A Preliminary Report from the International Council on Archives' Experts Group on Archival Description. *The American Archivist*, 76(2). <https://doi.org/10.17723/aarc.76.2.p071x02401282qx2>

Guia de classificació de documents del sistema general de gestió de la documentació administrativa. (2015). Departament de Cultura de la Generalitat de Catalunya.
https://cultura.gencat.cat/web/.content/dgpc/arxius_i_gestio_documental/14_Gestio_documento/02_GD_Generalitat/01_ATESTA/01_Protocol_i_Instruments/00_Instruments/01_quadre_classificacio_funcional/20150327_guia_QdCF.pdf

Guidance on a New Approach to Managing Email Records (Bulletin 2013-02). (2013). National Archives and Records Administration (NARA).
<https://www.archives.gov/records-mgmt/bulletins/2013/2013-02.html>

Gunnlaugsdottir, J. (2012a). Functional classification scheme for records: FCS: a way to chart documented knowledge in organizations. *Records Management Journal*, 22(2).
<https://doi.org/10.1108/09565691211268171>

Gunnlaugsdottir, J. (2012b). Information and records management: A precondition for a well functioning quality management system. *Records Management Journal*, 22(3).
<https://doi.org/10.1108/09565691211283138>

Hai, R., Geisler, S., & Quix, C. (2016). *Constance: An Intelligent Data Lake System*. SIGMOD/PODS'16: International Conference on Management of Data, San Francisco California USA. <https://doi.org/10.1145/2882903.2899389>

Hai, R., Koutras, C., Quix, C., & Jarke, M. (2021). Data Lakes: A Survey of Functions and Systems. *IEEE Transactions on Knowledge and Data Engineering*.
<https://doi.org/10.48550/ARXIV.2106.09592>

Halevy, A., Korn, F., Noy, N. F., Olston, C., Polyzotis, N., Roy, S., & Whang, S. E. (2016). *Goods: Organizing Google's Datasets*. SIGMOD/PODS'16: International Conference on Management of Data, San Francisco California USA.
<https://doi.org/10.1145/2882903.2903730>

- Harries, S. (2000). *Capturing and managing electronic records from websites and intranets in the government environment*. Proceedings of the DLM-Forum: European citizens and electronic information: the memory of the information society, Luxembourg. <https://www.project-consult.de/files/DLM%20Conference%202002.pdf>
- Harries, S. (2009). Managing records, making knowledge and good governance. *Records Management Journal*, 19(1). <https://doi.org/10.1108/09565690910937218>
- Harries, S. (2012). *Records management and knowledge mobilisation: A handbook for regulation, innovation and transformation*. Chandos Publishing Ltd.
- Hartig, O., & Zhao, J. (2010). Publishing and Consuming Provenance Metadata on the Web of Linked Data. En D. L. McGuinness, J. R. Michaelis, & L. Moreau (Ed.), *Provenance and Annotation of Data and Processes* (Vol. 6378). Springer International Publishing. https://doi.org/10.1007/978-3-642-17819-1_10
- Hawkins, A. (2021). *Advocating for Linked Archives: The benefits to users of Archival Linked Data*. Proceedings of Linked Archives International Workshop co-located with 25th International Conference on Theory and Practice of Digital Libraries (TPDL). <https://ceur-ws.org/Vol-3019/>
- Healy, S. (2001). ISO 15489 Records Management – its development and significance. *Records Management Journal*, 11. <https://doi.org/10.1108/EUM0000000007271>
- Hellberg, A. S., & Hedström, K. (2015). The story of the sixth myth of open data and open government. *Transforming Government: People, Process and Policy*, 9(1). <https://doi.org/10.1108/TG-04-2014-0013>
- Henttonen, P. (2009). A comparison of MoReq and SÄHKE metadata and functional requirements. *Records Management Journal*, 19(1). <https://doi.org/10.1108/09565690910937227>
- Henttonen, P., & Kettunen, K. (2011). Functional classification of records and organisational structure. *Records Management Journal*, 21(2). <https://doi.org/10.1108/09565691111152035>
- Heuscher, S., Jaermann, S., Keller-Marxer, P., & Moehle, F. (2004). Providing Authentic Long-term Archival Access to Complex Relational Data. *ArXiv*. <https://doi.org/10.48550/ARXIV.CS/0408054>
- Hilyer, L. A. (2014). Review: Archives and Recordkeeping: Theory into Practice. *The Journal of Academic Librarianship*, 40(3-4). <https://doi.org/10.1016/j.acalib.2014.05.010>
- Hofman, D., Lemieux, V. L., Joo, A., & Batista, D. A. (2019). “The margin between the edge of the world and infinite possibility”: Blockchain, GDPR and information governance. *Records Management Journal*, 29(1-2). <https://doi.org/10.1108/RMJ-12-2018-0045>

Hoke, G. E. J. (2022). Information Governance: The path from now to 2020. *Ethical Boardroom*. <http://ethicalboardroom.com/technology/information-governance-the-path-from-now-to-2020/>

Hurley, C. (2008). What, if anything, is the Australian "Series" System. Monash University. Journal contribution. <https://doi.org/10.4225/03/580576ede8e29>.

Ifould, P., & Joseph, P. (2016). User difficulties working with a business classification scheme: A case study. *Records Management Journal*, 26(1). <https://doi.org/10.1108/RMJ-03-2015-0011>

Iqbal, R., Yuda, P., Aditya, W., Hidayanto, A. N., Wuri Handayani, P., & Harahap, N. C. (2019). *Master Data Management Maturity Assessment: Case Study of XYZ Company*. 2nd International Conference on Applied Information Technology and Innovation (ICAITI), Denpasar, Bali, Indonesia. <https://doi.org/10.1109/ICAITI48442.2019.8982123>

ISO/TC46/SC11-Archives/records management. (2023). *Comparison of ISO 24143-Information Governance and ISO 30300 Management Systems for Records*. White paper. International Organization for Standardization. https://committee.iso.org/files/live/sites/tc46sc11/files/documents/Comparison%20of%20ISO%2024143%20and%20ISO%2030300%20series-Final_20230901.pdf

Jansen, A. (2015). Object-oriented diplomatics: Using archival diplomatics in software application development to support authenticity of digital records. *Records Management Journal*, 25(1). <https://doi.org/10.1108/RMJ-04-2014-0022>

Järvihaavisto, U., & Öhman, M. (2022). Building Trust in Business Ecosystems: The Interplay of Technology and Community in Governing Data Sharing. A D. Y. Kim, G. von Cieminski, & D. Romero (Ed.), *Advances in Production Management Systems. Smart Manufacturing and Logistics Systems: Turning Ideas into Action* (Vol. 663). Springer International Publishing. https://doi.org/10.1007/978-3-031-16407-1_43

Jenkinson, H. (1922). A Manual of archive administration: Including the problems of war archives and archive making. *A A Manual of archive administration: Including the problems of war archives and archive making*. Clarendon Press.

Johnson, T., & Dasu, T. (2003). *Data quality and data cleaning: An overview*. SIGMOD/PODS03: International Conference on Management of Data and Symposium on Principles Database and Systems, San Diego California. <https://doi.org/10.1145/872757.872875>

Jones, P. (2008). The role of virtual folders in developing an electronic document and records management system: Meeting user and records management needs. *Records Management Journal*, 18(1). <https://doi.org/10.1108/09565690810858514>

- Joseph, P., & Hartel, J. (2017). Visualizing information in the records and archives management (RAM) disciplines: Using Engelhardt's graphical framework. *Records Management Journal*, 27(3). <https://doi.org/10.1108/RMJ-06-2016-0017>
- Joyce, W. L. (2002). Enduring Paradigm, New Opportunities: The Value of the Archival Perspective in the Digital Environment (review). *Portal: Libraries and the Academy*, 2(1).
- Julbe, F. (2023). Evolució del processat de dades: Des dels processos manuals fins el Big Data i els nous paradigmes. *Lligall; Revista catalana d'arxivística*, 46. <https://raco.cat/index.php/liligall/article/view/416757>
- Kampffmeyer, U. (2000). E-documents: It's all legal, or is it? *E-doc*.
- Kampffmeyer, U. (2023). Governança de la informació: un manual. *Lligall; Revista catalana d'arxivística*, 46. <https://raco.cat/index.php/liligall/article/view/416756>
- Kandel, S., Heer, J., Plaisant, C., Kennedy, J., Van Ham, F., Riche, N. H., Weaver, C., Lee, B., Brodbeck, D., & Buono, P. (2011). Research directions in data wrangling: Visualizations and transformations for usable and credible data. *Information Visualization*, 10(4). <https://doi.org/10.1177/1473871611415994>
- Kaplan, R. S., & Norton, D. P. (1996). Linking the Balanced Scorecard to Strategy. *California Management Review*, 39(1). <https://doi.org/10.2307/41165876>
- Kastenhofer, J. (2015). The logic of archival authenticity: ISO 15489 and the varieties of forgeries in archives. *Archives and Manuscripts*, 43(3). <https://doi.org/10.1080/01576895.2015.1074085>
- Keng, S., & Xian, T. (2006). Use of Cognitive Mapping Techniques in Information Systems. *Journal of Database Management*, 17.
- Kent, W. (1983). A simple guide to five normal forms in relational database theory. *Communications of the ACM*, 26(2). <https://doi.org/10.1145/358024.358054>
- Kettunen, K., & Henttonen, P. (2010). Missing in action? Content of records management metadata in real life. *Library and Information Science Research*, 32(1). <https://doi.org/10.1016/j.lisr.2009.10.002>
- Khatri, V., & Brown, C. V. (2010). Designing data governance. *Communications of the ACM*, 53(1). <https://doi.org/10.1145/1629175.1629210>
- Khine, P. P., & Wang, Z. S. (2018). Data lake: A new ideology in big data era. *ITM Web of Conferences*, 17. <https://doi.org/10.1051/itmconf/20181703025>

Klischewski, R. (2003). *Towards an ontology for e-document management in public administration -the case of Schleswig-Holstein*. 36th Annual Hawaii International Conference on System Sciences. Proceedings of the IEEE, Hawaii.

Klischewski, R. (2006). Ontologies for e-document management in public administration. *Business Process Management Journal*, 12(1).
<https://doi.org/10.1108/14637150610643742>

Kononow, P. (2019, juliol 10). Business Glossary vs Data Glossary vs Data Dictionary. *Dataedo*. <https://dataedo.com/blog/business-glossary-vs-data-dictionary>

Kooper, M. N., Maes, R., & Lindgreen, E. E. O. R. (2011). On the governance of information: Introducing a new concept of governance to support the management of information. *International Journal of Information Management*, 31(3).
<https://doi.org/10.1016/j.ijinfomgt.2010.05.009>

Kudyba, S., Fjermestad, J., & Davenport, T. (2020). A research model for identifying factors that drive effective decision-making and the future of work. *Journal of Intellectual Capital*, 21(6). <https://doi.org/10.1108/JIC-05-2019-0130>

Kurth, M., Ruddy, D., & Rupp, N. (2004). Repurposing MARC metadata: Using digital project experience to develop a metadata management design. *Library Hi Tech*, 22(2).
<https://doi.org/10.1108/07378830410524585>

Kusuma, K. P., Fabrianti, T., & Andreswari, R. (2019). *Analysis and Design of Architecture Master Data Management (MDM) Tools for Open Source Platform at PT XYZ*. 5th International Conference on Science and Technology (ICST), Yogyakarta, Indonesia. <https://doi.org/10.1109/ICST47872.2019.9166255>

Labadie, C., Legner, C., Eurich, M., & Fadler, M. (2020). *FAIR Enough? Enhancing the Usage of Enterprise Data with Data Catalogs*. IEEE 22nd Conference on Business Informatics (CBI), Antwerp, Belgium. <https://doi.org/10.1109/CBI49978.2020.00029>

Ladley, J. (2016). Ensuring the quality of «fit for purpose» data. *CIO*.
<https://www.cio.com/article/236174/ensuring-the-quality-of-fit-for-purpose-data.html>

Langefors, B., & Dahlbom, B. (1995). *Essays on infology: Summing up and planning for the future*. Studentlitteratur.

Lappin, J. (2010). What will be the next records management orthodoxy? *Records Management Journal*, 20(3). <https://doi.org/10.1108/09565691011095283>

Lappin, J. (2020a, gener 30). Records management before and after the AI revolution. *Thinking Records*. <https://thinkingrecords.co.uk/>

- Lappin, J. (2020b, març 1). The records management implications of the rise of Microsoft Teams. *Thinking Records*. <https://thinkingrecords.co.uk/>
- Lappin, J. (2020c, març 22). Project Cortex and the future of document management in Office 365. *Thinking Records*. <https://thinkingrecords.co.uk/>
- Lappin, J. (2021, gener 1). The impact of Microsoft Teams on records management strategies. *Thinking Records*. <https://thinkingrecords.co.uk/>
- Lappin, J., Jackson, T., Matthews, G., & Ravenwood, C. (2021). Rival records management models in an era of partial automation. *Archival Science*, 21(3). <https://doi.org/10.1007/s10502-020-09354-9>
- Laurent, A., Laurent, D., & Madera, C. (Ed.). (2019). *Data lakes*. ISTE Ltd / John Wiley and Sons Inc.
- Lemieux, V. L. (2015). Visual analytics, cognition and archival arrangement and description: Studying archivists' cognitive tasks to leverage visual thinking for a sustainable archival future. *Archival Science*, 15(1). <https://doi.org/10.1007/s10502-013-9212-y>
- Lemieux, V. L. (2016). Trusting records: Is Blockchain technology the answer? *Records Management Journal*, 26(2). <https://doi.org/10.1108/RMJ-12-2015-0042>
- Lemieux, V. L., Gormly, B., & Rowledge, L. (2014). Meeting Big Data challenges with visual analytics: The role of records management. *Records Management Journal*, 24(2). <https://doi.org/10.1108/RMJ-01-2014-0009>
- Lemieux, V. L., Rowell, C., Seidel, M.-D. L., & Woo, C. C. (2020). Caught in the middle?: Strategic information governance disruptions in the era of blockchain and distributed trust. *Records Management Journal*, 30(3). <https://doi.org/10.1108/RMJ-09-2019-0048>
- Li, X., Yang, M., Xia, X., Zhang, K., & Liu, K. (2022). *A Distributed Data Fabric Architecture based on Metadata Knowledge Graph*. 5th International Conference on Data Science and Information Technology (DSIT), Shanghai, China. <https://doi.org/10.1109/DSIT55514.2022.9943831>
- Llanes-Padrón, D., & Pastor-Sánchez, J. A. (2017). Records in contexts: The road of archives to semantic interoperability. *Program: electronic library and information systems*, 51(4). <https://doi.org/10.1108/PROG-03-2017-0021>
- Lloveras, M. R. (2014). Aplicació documental de l'enfocament a processos. *Item: revista de biblioteconomia i documentació*, 1(58). <https://raco.cat/index.php/Item/article/view/284418>

Lo Giudice, P., Musarella, L., Sofo, G., & Ursino, D. (2019). An approach to extracting complex knowledge patterns among concepts belonging to structured, semi-structured and unstructured sources in a data lake. *Information Sciences*, 478.

<https://doi.org/10.1016/j.ins.2018.11.052>

Loshin, D. (2010). *Operationalizing Data Governance through Data Policy Management*. Knowledge Integrity, Inc.

Loshin, D. (2015). *Enterprise Data Architecture and Data Governance: Use Metadata to Get to the Starting Gate*. Knowledge Integrity, Inc.

Lourenço, A., Ramalho, J. C., Gago, M. R., & Penteado, P. (2020). Plataforma CLAV: Contributo para a disponibilização de dados abertos da Administração Pública em Portugal. *Cadernos BAD*, 19. <https://doi.org/10.48798/CADERNOSBAD.2047>

Ludlow, S. (2016). ECM Is Not Doomed to Fail. *Enterprise Information Management*. https://www.opentext.jp/file_source/OpenText/en_US/PDF/opentext-article-kmworld-stephen-ludlow-ecm-is-not-doomed-to-fail-en.pdf

Lueders, D. (2021, febrer 10). The Four Lies Destroying Records Management – Lie #1: Information vs Data. *Next Generation Records Management*. <https://nextgenrm.com/2021/02/10/the-four-lies-destroying-records-management-lie-1-information-vs-data/>

Macintosh, S., & Real, L. (2007). DIRKS: Putting ISO 15489 to Work. *Information Management Journal*, March/April.

Magić, I. (2008). MacNeil, H. Trusting Records: Legal, Historical and Diplomatic Perspectives. Dordrecht: Kulwer Academic Publishers, 2000. *Arhivski Vjesnik*, 51(1).

Makhlouf Shabou, B., Tièche, J., Knafou, J., & Gaudinat, A. (2020). Algorithmic methods to explore the automation of the appraisal of structured and unstructured digital data. *Records Management Journal*, 30(2). <https://doi.org/10.1108/RMJ-09-2019-0049>

Malmsten, M. (2013). Cataloguing in the open: The disintegration and distribution of the record. *JLIS.it*, 4(1).

Managing Government Records Directive (M-12-18). (2012). Office of Management and Budget (OMB) - National Archives and Records Administration (NARA).

Mandell, L., & Kriegsman, S. (2005). *Digital repository planning and policy*. Archiving 2005; Final Program and Proceedings, Springfield.

Maroye, L., vanHooland, S., ArangurenCelorrio, F., Soyez, S., Losdyck, B., Vanreck, O., & deTerwangne, C. (2017). Managing electronic records across organizational

boundaries: The experience of the Belgian federal government in automating investigation processes. *Records Management Journal*, 27(1).

<https://doi.org/10.1108/RMJ-11-2015-0037>

Marr, B. (2022, novembre 2). Beyond Dashboards: The Future Of Analytics And Business Intelligence? *Enterprise Tech - Forbes*.

<https://www.forbes.com/sites/bernardmarr/2022/11/02/beyond-dashboards-the-future-of-analytics-and-business-intelligence/?sh=294f18361804>

Marz, N., & Warren, J. (2015). *Big data: Principles and best practices of scalable real-time data systems*. Manning.

Mayer-Schönberger, V., & Cukier, K. (2014). *Big data: A revolution that will transform how we live, work, and think* (First Mariner Books edition). Mariner Books, Houghton Mifflin Harcourt.

Mazzini, S., & Ricci, F. (2011). *EAC-CPF Ontology and Linked Archival Data*. Proceedings of the 1st International Workshop on Semantic Digital Archives (SDA), Berlin. <https://ceur-ws.org/Vol-801/paper6.pdf>

McDonald, J. (1989). Records Management and Data Management: Closing the Gap. *Records Management Journal*, 1(1). <https://doi.org/10.1108/eb027016>

McDonald, J. (1995). Managing records in the modern office: Taming the wild frontier. *Archivaria*, 39.

McKemmish, S. (1997). *Yesterday, today and tomorrow: A continuum of responsibility*. Proceedings of the Records Management Association of Australia 14th National Convention, Perth. <https://doi.org/10.4225/03/580576155de57>

McKemmish, S., Acland, G., Ward, N., & Reed, B. (1999). Describing Records in Context in the Continuum: The Australian Recordkeeping Metadata Schema. *Archivaria*, 48.

McKemmish, S., Chandler, T., & Faulkhead, S. (2019). Imagine: A living archive of people and place "somewhere beyond custody". *Archival Science*, 19(3).

<https://doi.org/10.1007/s10502-019-09320-0>

McLeod, J. (2008). Understanding Data and Information Systems for Recordkeeping. *Records Management Journal*, 18(2). <https://doi.org/10.1108/rmj.2008.28118bae.003>

Méndez, E. Ma. (2003). Catalogació/organització de documents digitals: Estat de la qüestió, tendències i perspectives des d'Espanya. *Bibliodoc: anuari de biblioteconomia, documentació i informació*.

- Menne-Haritz, A. (2001). Access—The reformulation of an archival paradigm. *Archival Science*, 1(1). <https://doi.org/10.1007/BF02435639>
- Millar, L. (2017). On the crest of a wave: Transforming the archival future. *Archives and Manuscripts*, 45(2). <https://doi.org/10.1080/01576895.2017.1328696>
- Milne, C. (2007). Taxonomy development: Assessing the merits of contextual classification. *Records Management Journal*, 17(1). <https://doi.org/10.1108/09565690710730660>
- Mokhtar, U. A., & Yusof, Z. M. (2015). Classification: The understudied concept. *International Journal of Information Management*, 35(2). <https://doi.org/10.1016/j.ijinfomgt.2014.12.002>
- Mokhtar, U. A., & Yusof, Z. M. (2016). Records management practice: The issues and models for classification. *International Journal of Information Management*, 36(6). <https://doi.org/10.1016/j.ijinfomgt.2016.09.001>
- Mokhtar, U. A., Yusof, Z. M., Ahmad, K., & Jambari, D. I. (2016). Development of function-based classification model for electronic records. *International Journal of Information Management*, 36(4). <https://doi.org/10.1016/j.ijinfomgt.2016.04.009>
- Moon, S.-J., Kang, S.-B., & Park, B.-J. (2021). A Study on a Distributed Data Fabric-based Platform in a Multi-Cloud Environment. *International Journal of Advanced Culture Technology (IJACT)*, 9(3).
- Mordell, D. (2019). Critical Questions for Archives as (Big) Data. *Archivaria*, 87.
- Mordell, D. (2022). *Oil, abundance, and intangibility: Mapping an (archives as) data paradigm*. Digital Archives, Big Data and Memory, Copenhagen.
- Morelli, J. (2007). Hybrid filing schemes: The use of metadata signposts in functional file plans. *Records Management Journal*, 17(1). <https://doi.org/10.1108/09565690710730679>
- Moss, M., Thomas, D., & Gollins, T. (2018). The Reconfiguration of the Archive as Data to Be Mined. *Archivaria*, 86.
- Moyano, J. (2013). La descripción archivística. De los instrumentos de descripción hacia la web semántica. *Anales de Documentación*, 16(2). <https://doi.org/10.6018/analesdoc.16.2.171841>
- Munn, L. (2022). The uselessness of AI ethics. *AI and Ethics*, 3. <https://doi.org/10.1007/s43681-022-00209-w>

- Mutschler, B., Weber, B., & Reichert, M. (2008). *Workflow management versus case handling: Results from a controlled software experiment*. SAC '08: The ACM Symposium on Applied Computing, Fortaleza, Brazil. <https://doi.org/10.1145/1363686.1363711>
- Niu, J. (2013). Recordkeeping metadata and archival description: A revisit. *Archives and Manuscripts*, 41(3). <https://doi.org/10.1080/01576895.2013.829751>
- Niu, J. (2014). Event-based archival information organization. *Archival Science*, 15(3). <https://doi.org/10.1007/s10502-014-9222-4>
- Niu, J. (2015). Original order in the digital world. *Archives and Manuscripts*, 43(1). <https://doi.org/10.1080/01576895.2014.958863>
- Niu, J. (2016). Organisation and description of datasets. *Archives and Manuscripts*, 44(2). <https://doi.org/10.1080/01576895.2016.1179585>
- Ojala, M. (2021). Understanding Enterprise Information. *KM World*, 30(6).
- Oladejo, B., & Hadžidedić, S. (2021). Electronic records management – a state of the art review. *Records Management Journal*, 31(1). <https://doi.org/10.1108/RMJ-09-2019-0059>
- Orr, S. (2006). Functions-based classification of records: Is it functional? *Archives and Manuscripts*, 34(1).
- Östberg, P.-O., Vyhmeister, E., Castañé, G. G., Meyers, B., & Van Noten, J. (2022). Domain Models and Data Modeling as Drivers for Data Management: The ASSISTANT Data Fabric Approach. *IFAC PapersOnLine*, 55(10).
- Packalén, S. (2015). Functional classification. *Records Management Journal*, 25(2). <https://doi.org/10.1108/RMJ-10-2014-0043>
- Padilla, T. (2017). *On a Collections as Data Imperative*. https://digitalpreservation.gov/meetings/dcs16/tpadilla_OnaCollectionsasDataImperative_final.pdf
- Pan, W. (2017). The implementation of electronic recordkeeping systems: An exploratory study of socio-technical issues. *Records Management Journal*, 27(1). <https://doi.org/10.1108/RMJ-04-2016-0014>
- Papadakis, G., Alexiou, G., Papastefanatos, G., & Koutrika, G. (2015). Schema-agnostic vs schema-based configurations for blocking methods on homogeneous data. *Proceedings of the VLDB Endowment*, 9(4). <https://doi.org/10.14778/2856318.2856326>

Papadakis, G., Papastefanatos, G., & Koutrika, G. (2014). Supervised meta-blocking. *Proceedings of the VLDB Endowment*, 7(14).

<https://doi.org/10.14778/2733085.2733098>

Papastefanatos, G., Meimaris, M., & Vassiliadis, P. (2022). Relational schema optimization for RDF-based knowledge graphs. *Information Systems*, 104.

<https://doi.org/10.1016/j.is.2021.101754>

Parilla, L., Morgan, R., & Fidler, C. (2017). Excavating archival description: From collection to data level. *Digital Library Perspectives*, 33(3). <https://doi.org/10.1108/DLP-11-2016-0043>

Parycek, P., Höchtl, J., & Ginner, M. (2014). Open government data implementation evaluation. *Journal of Theoretical and Applied Electronic Commerce Research*, 9(2).

<https://doi.org/10.4067/S0718-18762014000200007>

Patton, S. (2019, novembre 4). Introducing Project Cortex. *Microsoft 365 Blog*.

<https://techcommunity.microsoft.com/t5/microsoft-365-blog/introducing-project-cortex/ba-p/966091>

Payne, N. (2019). *An Intelligent Class: The Development Of A Novel Context Capturing Framework Supporting The Functional Auto-Classification Of Records*. IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA.

<https://doi.org/10.1109/BigData47090.2019.9006564>

Payne, N. (2022). *An Intelligent Class: The Development Of A Novel Context Capturing Method For The Functional Auto Classification Of Records*. IEEE International Conference on Big Data (Big Data), Osaka, Japan.

<https://doi.org/10.1109/BigData55660.2022.10020387>

Penteado, P. (2015). Gestão de documentos de arquivo na administração pública em Portugal: Experiência e desafios. *Acervo*, 28(2).

Penteado, P. (2018). Políticas e práticas de gestão de documentos eletrônicos na Administração: O que está mudando em Portugal. A *El archivo electrónico en la administración digital: 23 Jornadas de Archivos Universitarios, 21-23 de junio de 2017*. Universidade da Coruña - Servizo de Publicacións.

<https://doi.org/10.17979/spudc.9788497496803.061>

Pérez-Montoro, M. (2020). Aportaciones en el campo de la visualización de información en 2019. *Anuario ThinkEPI*, 14.

<https://doi.org/10.3145/thinkepi.2020.e14d01>

Pires, R. (2023, março 20). AI Is Coming To Microsoft 365: How Copilot Will Change The Way You Use Your Favorite Tools. *Rockcontent*.

<https://rockcontent.com/blog/microsoft-365-copilot/>

Puig-Pey, A. de P., & Guiu, P. (2002). Auditoria de circuits administratius: Anàlisi i propostes de millora. En *Auditoria de circuits administratius: Anàlisi i propostes de millora*. Associació d'Arxivers Valencians.

Puig-Pey, A. de P., & Guiu, P. (2008). Circuits administratius: Disseny i millora. En *Circuits administratius: Disseny i millora*. Associació d'Arxivers Valencians.

Puig-Pey, A. de P., & Guiu, P.-J. (2013). Simplificació vers l'e-Administració. Redisseny i millora d'expedients per unificar procediments, estandarditzar documentació i reduir continguts d'informació. *Estudis de Recerca Digitals*, 10. <https://doi.org/10.2436/10.8030.03.6>

Pyzdek, T., & Keller, P. (2018). *The Six sigma handbook: A complete guide for green belts, black belts, and managers at all levels* (Fifth edition). McGraw-Hill Education.

Quílez, J. (2011). Aproximació a les ontologies: Definició i construcció. Aplicació al camp de l'arxivística. *Lligall; Revista catalana d'arxivística*, 32. <https://raco.cat/index.php/liligall/article/view/340074>

Quílez, J. (2017). Integració d'arxius, museus i biblioteques amb tecnologies del web semàntic: De la modelització conceptual a la iniciativa linked open data. *Lligall; Revista catalana d'arxivística*, 40. <https://raco.cat/index.php/liligall/article/view/340577>

Quimbert, E., Jeffery, K., Martens, C., Martin, P., & Zhao, Z. (2020). Data cataloguing. En Z. Zhao & M. Hellström (Ed.), *Towards Interoperable Research Infrastructures for Environmental and Earth Sciences: A Reference Model Guided Approach for Common Challenges* (Vol. 12003). Springer International Publishing. <https://doi.org/10.1007/978-3-030-52829-4>

Quix, C., Hai, R., & Vatov, I. (2016). *GEMMS: A Generic and Extensible Metadata Management System for Data Lakes*. Proceedings of the CAiSE'16 Forum at the 28th International Conference on Advanced Information Systems Engineering, Ljubljana, Slovenia. <http://ceur-ws.org>

Rajmohan, Chaudhuri, R., Ganesan, B., Sanghi, A., Agarwal, A., & Mehta, S. (2023). *Tutorial on Semantic Automation for Data Discovery*. CODS-COMAD: 6th Joint International Conference on Data Science & Management of Data (10th ACM IKDD CODS and 28th COMAD), Mumbai India. <https://doi.org/10.1145/3570991.3571061>

Ramalho, José Carlos, Ferreira, Miguel, Faria, Luís, & Castro, Rui. (2007). *Relational database preservation through XML modelling*. Extreme Markup Languages Conference, Montreal. <https://hdl.handle.net/1822/7120>

Rasmussen, K. B., & Blank, G. (2007). The data documentation initiative: A preservation standard for research. *Archival Science*, 7(1).

- Redkina, N. S. (2019). Current Trends in Research Data Management. *Scientific and Technical Information Processing*, 46(2). <https://doi.org/10.3103/S0147688219020035>
- Reina, L. (2023). Noves arquitectures de dades. *Lligall; Revista catalana d'arxivística*, 46. <https://raco.cat/index.php/lligall/article/view/416761>
- Rhoads, J. (1969). The Historian and the New Technology. *The American Archivist*, 32(3). <https://doi.org/10.17723/aarc.32.3.f8181nm6mjv147t7>
- Richards, L. L. (2016). Teaching data creators how to develop an OAIS-compliant digital curation system: Colearning and breakdowns in support of requirements analysis. *American Archivist*, 79(2). <https://doi.org/10.17723/0360-9081-79.2.371>
- Riley, J. (2017). Understanding metadata: What is metadata, and what is it for? En *Understanding metadata: What is metadata, and what is it for?* National Information Standards Organization NISO.
- Ritter, J., Jenkins, J., & Pratt, M. K. (2015). *Data Governance strategies for the digital Age*. TechTarget.
- Rittgen, P. (2006). A language-mapping approach to action-oriented development of information systems. *European Journal of Information Systems*, 15(1). <https://doi.org/10.1057/palgrave.ejis.3000597>
- Roberge, M. (1992). *La Gestió dels documents administratius*. Associació d'Arxivers de Catalunya.
- Rodríguez, J. M. (2012). Series y subseries en un sistema de clasificación funcional. *Anales de Documentación*, 15(2). <https://doi.org/10.6018/analesdoc.15.2.147931>
- Rogers, C. (2015). Diplomatics of born digital documents – considering documentary form in a digital environment. *Records Management Journal*, 25(1). <https://doi.org/10.1108/RMJ-03-2014-0021>
- Rolan, G., Humphries, G., Jeffrey, L., Samaras, E., Antsoupova, T., & Stuart, K. (2019). More human than human? Artificial intelligence in the archive. *Archives and Manuscripts*, 47(2). <https://doi.org/10.1080/01576895.2018.1502088>
- Rosing, M. V. (2014). *The complete business process handbook: Body of knowledge from process modeling to bpm* (1st edition). Elsevier.
- Sabourin, P. (2001). Constructing a Function-Based Records Classification System: Business Activity Structure Classification System. *Archivaria*, 51.
- Saffady, W. (2014). Taking control of e-mail with uniform retention rules. *Information Management Journal*, 48(1). Gale Academic OneFile.

- Sarkar, P. (2015). *Data as a service: A framework for providing reusable enterprise data services*. IEEE.
- Sataslaatten, O. H. (2014). The Norwegian Noark Model requirements for EDRMS in the context of open government and access to governmental information. *Records Management Journal*, 24(3). <https://doi.org/10.1108/RMJ-09-2014-0041>
- Sawadogo, P., & Darmont, J. (2021). On data lake architectures and metadata management. *Journal of Intelligent Information Systems*, 56(1). <https://doi.org/10.1007/s10844-020-00608-7>
- Schellenberg, Theodore R. (1956). *Modern archives: Principles and techniques*. University of Chicago Press.
- Schellenberg, Theodore R. (1965). *The Management of Archives*. Columbia University Press.
- Schmid, S., Henson, C., & Tran, T. (2019). Using Knowledge Graphs to Search an Enterprise Data Lake. A P. Hitzler, S. Kirrane, O. Hartig, V. de Boer, M.-E. Vidal, M. Maleshkova, S. Schlobach, K. Hammar, N. Lasierra, S. Stadtmüller, K. Hose, & R. Verborgh (Ed.), *The Semantic Web: ESWC 2019 Satellite Events* (Vol. 11762). Springer International Publishing. https://doi.org/10.1007/978-3-030-32327-1_46
- Schwalm, S., & Alamillo, I. (2021). Self-Sovereign-Identity & eIDAS: a Contradiction? Challenges and Chances of eIDAS 2.0. *European Review of Digital Administration & Law*, 2(2).
- Scott, P. (1966). The Record Group Concept: A Case for Abandonment. *The American Archivist*, 29(4). <https://doi.org/10.17723/aarc.29.4.y886054240174401>
- Serifi, V., Dašić, P., Ječmenica, R., & D.Labović. (2013). Functional and information modeling of production using IDEF methods. *Strojnicki Vestnik*, 55.
- Serra, J. (2006). La gestió per expedients i la preservació de la documentació electrònica: El sistema BAULA. *Lligall; Revista catalana d'arxivística*, 25. <https://raco.cat/index.php/lligall/article/view/339633>
- Serra, J. (2010, maig 7). La inclusión de requisitos archivísticos en proyectos de tramitación electrónica: El mapa de series y sus formas de implantación. *Revista d'arxius*, 8.
- Serra, J. (2013). *Una interpretació metodològica de la norma ISO 15489 para la implantación de un sistema de gestión de documentos*. Jornadas Ibéricas de Arquivos Municipais; Políticas, Sistemas e Instrumentos, Lisboa.

Serra, J. (2021). Los datos: ¿una amenaza? : Gestión de documentos en la era de los datos. *Tabula: revista de archivos de Castilla y León*, 24.

Serra, J., & Casellas, M. (2007). Un model d'implantació d'un sistema de gestió de la documentació administrativa: El projecte DursiGED. *BiD: textos universitaris de biblioteconomia i documentació*, 18. <https://raco.cat/index.php/BiD/article/view/72767>

Shanmugam, S., & Seshadri, G. (2016). *Aspects of Data Cataloguing for Enterprise Data Platforms*. IEEE 2nd International Conference on Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing (HPSC) and IEEE International Conference on Intelligent Data and Security (IDS), New York, NY, USA. <https://doi.org/10.1109/BigDataSecurity-HPSC-IDS.2016.52>

Shepherd, E. (2006). Why are records in the public sector organizational assets? *Records Management Journal*, 16(1). <https://doi.org/10.1108/09565690610654747>

Shepherd, E. (2014). *What value does records management add to government administrative data?* https://interparestrust.org/assets/public/dissemination/ICA_shepherd_short2014%5B1%5D.pdf

Shepherd, E., Bunn, J., Flinn, A., Lomas, E., Sexton, A., Brimble, S., Chorley, K., Harrison, E., Lowry, J., & Page, J. (2019). Open government data: Critical information management perspectives. *Records Management Journal*, 29(1-2). <https://doi.org/10.1108/RMJ-08-2018-0023>

Shepherd, E., & Smith, C. (2000). The Application of ISAD(G) to the Description of Archival Datasets. *Journal of the Society of Archivists*, 21(1). <https://doi.org/10.1080/00379810050006911>

Shepherd, E. J., & Yeo, G. (2003). *Managing records: A handbook of principles and practice*. Facet.

Sherman, S. (2021, novembre 18). Retrospective Governance in M365. *Leadership through data*. <https://www.leadershipthroughdata.com.au/latest-news/retrospective-governance-in-m365/>

Sierra, L. F. (2012). Gestión Documental enfocada a procesos: Una mirada desde la administración pública distrital. *Revista Interamericana de Bibliotecología*, 35(3).

Skjekkeland, A. (2021, abril 5). From Big Buckets to Records Management in M365. *Infotechtion*. <https://www.infotechtion.com/post/from-big-buckets-to-records-management-in-m365>

Smajevic, M., & Bork, D. (2021). *From Conceptual Models to Knowledge Graphs: A Generic Model Transformation Platform*. ACM/IEEE International Conference on Model

Driven Engineering Languages and Systems Companion (MODELS-C), Fukuoka, Japan. <https://doi.org/10.1109/MODELS-C53483.2021.00093>

Smajevic, M., Hacks, S., & Bork, D. (2021). Using Knowledge Graphs to Detect Enterprise Architecture Smells. A E. Serral, J. Stirna, J. Ralyté, & J. Grabis (Ed.), *The Practice of Enterprise Modeling* (Vol. 432). Springer International Publishing. https://doi.org/10.1007/978-3-030-91279-6_4

Smit, F., Glaudemans, A., & Jonker, R. (Ed.). (2017). *Archives in liquid times*. Stichting Archiefpublicaties.

Soler, J. (2011). Una diplomàtica pels documents en crisi. *Lligall; Revista catalana d'arxivística*, 32. <https://raco.cat/index.php/liligall/article/view/340072>

Soler, J. (2020, setembre 29). Governança de dades, gestió documental o governança de la informació? *EAPC blog*. https://eapc.blog.gencat.cat/2020/09/29/governanca-de-dades-gestio-documental-o-governanca-de-la-informacio-joan-soler-jimenez/?utm_source=twitter&utm_medium=social&utm_campaign=

Spataro, J. (2023, març 16). Introducing Microsoft 365 Copilot – your copilot for work. *Official Microsoft Blog*. <https://blogs.microsoft.com/blog/2023/03/16/introducing-microsoft-365-copilot-your-copilot-for-work/>

Stamper, J., & Cervellati, G. (2023). *How Data Virtualization Offers a Road Map to Enterprise Sustainability*. IDC - Denodo.

Stillerman, J., Fredian, T., Greenwald, M., & Manduchi, G. (2016). Data catalog project—A browsable, searchable, metadata system. *Fusion Engineering and Design*, 112(C).

Stine, K., Kissel, R., Barker, W. C., & Fahlsing, J. (2008). Guide for Mapping Types of Information and Information Systems to Security Categories. *Network Security*, 1.

Suderman, J. (1997). Appraising Records of the Expenditure Management Function: An Exercise in Functional Analysis. *Archivaria*, 43.

Suderman, J. (2002). Defining Electronic Series: A Study. *Archivaria*, 53.

Sundberg, H. (2013). Process based archival descriptions – organizational and process challenges. *Business Process Management Journal*, 19(5). <https://doi.org/10.1108/BPMJ-Jan-2012-0002>

Sundgren, B. (2003). *Exploring patterns in information management: Concepts and perspectives for understanding IT-related change*. Stockholm School of Economics - EFI, the Economic Research Institute.

- Svärd, P. (2014). The impact of information culture on information/records management. *Records Management Journal*, 24(1). <https://doi.org/10.1108/RMJ-04-2013-0007>
- Térmens, Miquel. (2013). Preservación digital. A *Preservación digital* (Vol. 16). Universitat Oberta de Catalunya.
- Thomson, S. D. (2016). *Preserving Transactional Data* (Núm. 16; DPC Technology Watch Report). Digital Preservation Coalition. <https://doi.org/10.7207/twr16-02>
- Todd, M. (2003). *Business classification scheme design* (1.0). The National Archives.
- Torreblanca, S. (2023). La governança de dades com a interacció: Un concepte analític per a les administracions públiques. *Lligall; Revista catalana d'arxivística*, 46. <https://raco.cat/index.php/lligall/article/view/416759>
- Upward, F. (1996). Structuring the records continuum (Series of two parts) Part 1: Post custodial principles and properties. *Archives and Manuscripts*, 24(2).
- Upward, F. (1997). Structuring the records continuum (Series of two parts) Part 2: Structuration theory and recordkeeping. *Archives and Manuscripts*, 25(1).
- Upward, F. (2000). Modelling the continuum as paradigm shift in recordkeeping and archiving processes, and beyond Ö a personal reflection. *Records Management Journal*, 10(3). <https://doi.org/10.1108/EUM0000000007259>
- Upward, F. (2019). The monistic diversity of continuum informatics: A method for analysing the relationships between recordkeeping informatics, ethics and information governance. *Records Management Journal*, 29(1-2). <https://doi.org/10.1108/RMJ-09-2018-0028>
- Upward, F., Reed, B., Oliver, G., & Evans, J. (2013). Recordkeeping informatics: Re-figuring a discipline in crisis with a single minded approach. *Records Management Journal*, 23(1). <https://doi.org/10.1108/09565691311325013>
- Ur, A., Muzammal, M., David, G., & Ribeiro, C. (2015). Database Preservation: The DBPreserve Approach. *International Journal of Advanced Computer Science and Applications*, 6(12). <https://doi.org/10.14569/IJACSA.2015.061235>
- Van Der Aalst, W. M. P., Ter Hofstede, A. H. M., & Weske, M. (2003). Business Process Management: A Survey. A W. M. P. Van Der Aalst & M. Weske (Ed.), *Business Process Management* (Vol. 2678). Springer International Publishing. https://doi.org/10.1007/3-540-44895-0_1

Van Der Aalst, W. M. P., Weske, M., & Grünbauer, D. (2005). Case handling: A new paradigm for business process support. *Data & Knowledge Engineering*, 53(2). <https://doi.org/10.1016/j.datak.2004.07.003>

Van Dijck, J. (2014). Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & Society*, 12(2). <https://doi.org/10.24908/ss.v12i2.4776>

Vardigan, M., Heus, P., & Thomas, W. (2008). Data Documentation Initiative: Toward a Standard for the Social Sciences. *International Journal of Digital Curation*, 3(1).

Vázquez Murillo, M. (1995). *Manual de selección documental*. S&C.

Vázquez Murillo, M. (2006). *Cómo seleccionar documentos de archivo*. Alfagrama.

Vicknair, C. (2010). *Research issues in data provenance*. The 48th Annual Southeast Regional Conference - ACM SE '10, Oxford, Mississippi. <https://doi.org/10.1145/1900008.1900037>

Vieira, R., Ferreira, D., Borbinha, J., & Gaspar, G. (2012). A requirements engineering analysis of MoReq. *Records Management Journal*, 22(3).

Waddington, W., McGinley, K., Chu, P. K. J., Georgievski, G., & Kulkarni, D. (2021). *Architecting Data-Intensive SaaS Applications*. O'Reilly Media, Inc.

Walker, C. (2023, març 20). *Principles of Holistic Information Governance – An Introduction*. Digital Clarity Group. <http://www.digitalclaritygroup.com/phigs-introduction/>

Wang, X. (2014). *An analysis of the benefits and issues in the development of an Enterprise Data Catalogue*. Victoria University.

White Paper on The Capstone Approach and Capstone GRS. (2015). National Archives and Records Administration (NARA). <https://www.archives.gov/files/records-mgmt/email-management/final-capstone-white-paper.pdf>

Wieder, P., & Nolte, H. (2022). Toward data lakes as central building blocks for data management and analysis. *Frontiers in Big Data*, 5. <https://doi.org/10.3389/fdata.2022.945720>

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Mons, B. (2016). Comment: The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3. <https://doi.org/10.1038/sdata.2016.18>

Williams, R. D., Shankar, K., & Eschenfelder, K. R. (2017). Two views of the data documentation initiative: Stakeholders, collaboration and metadata standards creation. *Proceedings of the ASIST Annual Meeting*, 54(1).

Wilson, T. (2017). Rethinking digital preservation: Definitions, models, and requirements. *Digital Library Perspectives*, 33(2). <https://doi.org/10.1108/DLP-08-2016-0029>

Wright, T. (2013). Information culture in a government organization: Examining records management training and self-perceived competencies in compliance with a records management program. *Records Management Journal*, 23(1). <https://doi.org/10.1108/09565691311325004>

Yakel, E. (2000). Thinking Inside and Outside the Boxes: Archival Reference Services at the Turn of the Century. *Archivaria*, 49.

Yamaoka, E. J., & Gauthier, F. O. (2013). Objetos digitais: Em busca da precisão conceitual. *Informação & Informação*, 18(2). <https://doi.org/10.5433/1981-8920.2013v18n2p77>

Yeo, G. (2007). Concepts of Record (1): Evidence, Information, and Persistent Representations. *The American Archivist*, 70(2). <https://doi.org/10.17723/aarc.70.2.u327764v1036756q>

Yeo, G. (2009). Introduction to the series. En *Community archives: The shaping of memory*. Facet.

Yeo, G. (2011). Rising to the level of a record? Some thoughts on records and documents. *Records Management Journal*, 21(1). <https://doi.org/10.1108/09565691111125071>

Yeo, G. (2012). Bringing Things Together: Aggregate Records in a Digital Age. *Archivaria*, 74.

Yeo, G. (2013). Archival description in the era of digital abundance. *Comma*, 2. <https://doi.org/10.3828/comma.2013.2.2>

Zaidi, E., Guido, D. S., Roxane, E., & Alan, D. D. (2017). *Data catalogs are the new black in data management and analytics* (Núm. G00338777). Gartner. <https://www.gartner.com/en/documents/3837968/data-catalogs-are-the-new-black-in-data-management-and-a>

Zapata, C. A. (2002). *Mapas Documentales: Concepto y utilidad* (Gestión de información empresarial). http://eprints.rclis.org/9895/1/Mapas_Documentales.pdf

Zhang, J. (2015). Correspondence as a documentary form, its persistent representation, and email management, preservation, and access. *Records Management Journal*, 25(1). <https://doi.org/10.1108/RMJ-03-2014-0015>

Zikopoulos, P., Roos, D. de, Bienko, C., Buglio, R., & Andrews, M. (2015). *Big data beyond the hype: A guide to conversations for today's data center*. Mc Graw Hill Education.

