

Tesis doctoral

Ontología de características de la baldosa cerámica desde la Terminología

Autora: Ana María Estellés Palanca

Directora: Amparo Alcina Caudet

Universitat Jaume I

Diciembre de 2013



A mi familia y, en particular, a Eloy.

Agradecimientos

Agradecimientos a Amparo Alcina por el tiempo y el esfuerzo dedicado en la dirección de esta tesis. Doy las gracias también a todos mis compañeros del Departamento de Traducción, especialmente a Rosa Agost, porque sin su apoyo y consejo no habría terminado esta tesis. Me gustaría dar las gracias también a mis compañeros de Luxemburgo, especialmente a María José Palos, por todo lo que me han enseñado y por favorecer el mejor de los climas de trabajo.

No me puedo olvidar de dar las gracias a Isaac Nebot Díaz por sus valiosas aportaciones como experto en cerámica y por el interés mostrado en este proyecto, al Instituto de Tecnología Cerámica por asesorarme y facilitar el acceso a sus recursos y a las distintas empresas cerámicas que me facilitaron sus catálogos y me dieron acceso a numerosa información. Me gustaría reconocer especialmente el trabajo y el interés mostrado por Ernesto Jiménez y Tanguy Wettengel en la revisión de esta tesis.

A lo largo de esta tesis he tenido la oportunidad de hacer grandes amigos como Marta, Esther, Anna, Nuria, Dora, Gloria, Pilar, Silvia, Vero, Espe, José Luis, Ximo y muchos más. A todos: ¡muchas gracias!

Finalmente, me gustaría agradecer muy especialmente la comprensión, la paciencia y el ánimo recibidos de mi familia y amigos, ha sido una suerte.

Tesis doctoral

Ontología de características de la baldosa cerámica desde la Terminología

Autora: Ana María Estellés Palanca

Directora: Amparo Alcina Caudet

Universitat Jaume I

Diciembre de 2013



Desde distintos puntos de vista, dos hombres miran el mismo paisaje. Sin embargo, no ven lo mismo. La distinta situación hace que el paisaje se organice entre ambos de distinta manera. ¿Tendría sentido que cada uno declarara falso el paisaje ajeno? Evidentemente, no; tan real es el uno como el otro. La realidad cósmica es tal que solo puede ser vista bajo una determinada perspectiva. La perspectiva es uno de los componentes de la realidad. Lejos de ser su deformación, es su organización. Una realidad que vista desde cualquier punto resultara siempre idéntica es un concepto absurdo.”

(J. Ortega y Gasset)

Índice de contenidos

I.	INTRODUCTION	11
1	OVERVIEW	12
2	ASSUMPTIONS AND GOALS	14
3	METHODOLOGY	14
4	STRUCTURE	15
II.	FUNDAMENTOS TEÓRICOS	18
1	ENFOQUES EN EL ESTUDIO DE LA TERMINOLOGÍA	19
1.1	La Teoría General de la Terminología vs la interdisciplinariedad de la terminología	19
1.2	Representación de conocimiento en Terminología según Meyer	21
1.3	Terminología Socio-Cognitiva	24
1.4	Terminología Basada en Marcos	27
1.5	Ontoterminología	29
1.6	Recapitulación y observaciones	32
2	LAS CARACTERÍSTICAS	34
2.1	El término	34
2.1.1	El concepto	35
2.1.2	La intensión y extensión de un concepto	36
2.1.3	Tipos de concepto	37
2.1.4	Definición de característica	38
2.1.5	Recapitulación y observaciones	41
2.2	Tipos y partes de características	42
2.2.1	Características esenciales, distintivas y no esenciales	42
2.2.2	Nombres de característica y valores de características	44
2.2.3	Características intrínsecas y extrínsecas	44
2.2.4	Características como categorías descriptivas	47
2.2.5	Características según el ámbito	49
2.2.6	Características equivalentes	50
2.2.7	Recapitulación y observaciones	51

2.3	Relaciones de las características	52
2.3.1	Propiedades matemáticas de las relaciones conceptuales	53
2.3.2	Relaciones lógicas	54
2.3.3	Relaciones parte-todo	54
2.3.4	Relaciones secuenciales	56
2.3.5	Relaciones argumentales y circunstanciales	56
2.3.6	Otras relaciones	57
2.3.7	Recapitulación y observaciones	58
3	REPRESENTACIÓN DE CONCEPTOS EN TERMINOLOGÍA	59
3.1	Las definiciones terminológicas	59
3.2	Las clasificaciones jerárquicas	60
3.3	Los tesauros	61
3.4	Los bancos de datos terminológicos	62
3.5	Bases de conocimiento terminológico y sistemas orientados a ontologías	63
3.5.1	Cogniterm	64
3.5.2	CAOS (Computer-Aided Ontology Structuring)	66
3.5.3	GENOMA-KB	67
3.5.4	EcoLexicon	69
3.5.5	Oncoterm	72
3.6	Recapitulación y observaciones	73
4	METODOLOGÍAS PARA LA CREACIÓN DE RECURSOS TERMINOLÓGICOS	75
4.1	Metodología para la elaboración de un banco de datos terminológico	75
4.2	Metodologías sobre elaboración de bases de conocimiento terminológico y ontologías	75
4.2.1	Metodología para elaborar una base de conocimiento terminológico	76
4.2.2	Metodología de la Teoría Basada en Marcos	78
4.2.3	Metodología para transformar un modelo de conceptos en un modelo de datos conceptuales	81
4.2.4	Metodología Termontography	83
4.2.5	Metodología de la Ontoterminología	85
4.3	Recapitulación y observaciones	86
5	HERRAMIENTAS PARA LA REPRESENTACIÓN DE CONCEPTOS EN TERMINOLOGÍA	89
5.1	CODE4	89
5.2	Ontoterm	91

5.3	Termonography Workbench	93
5.3.1	Multilingual Categorization Framework Editor	95
5.4	Recapitulación y observaciones	97
6	INGENIERÍA DEL CONOCIMIENTO: LAS ONTOLOGÍAS	98
6.1	El concepto de ‘característica’ en Ingeniería del Conocimiento	98
6.1.1	Cualidades y propiedades	98
6.1.2	Cualidades y <i>qualia</i>	98
6.1.3	Propiedades y características	99
6.1.4	Recapitulación y observaciones	100
6.2	Representación de conceptos en Ingeniería del Conocimiento	100
6.2.1	Esquemas de representación formal	102
6.2.2	Lógica de primer orden	102
6.2.3	Redes semánticas	102
6.2.4	Marcos	104
6.2.5	La lógica de descripciones	105
6.2.6	Recapitulación y observaciones	107
6.3	Ontologías	107
6.3.1	Ejemplos de ontologías	110
6.3.2	Recapitulación y observaciones	114
7	EL LENGUAJE ESTÁNDAR DE ONTOLOGÍAS OWL	116
7.1	Componentes del lenguaje OWL 2	116
7.1.1	Entidades OWL 2	117
7.1.2	Descripción de tipos de clases e instancias OWL 2	121
7.1.3	Descripción de tipos de propiedades OWL 2	123
7.1.4	Descripción de conceptos a través de clases	126
7.2	Manchester OWL 2 Syntax	128
7.2.1	Restricciones	128
7.2.2	Construcciones de clases booleanas	129
7.2.3	Expresiones de clases complejas	129
7.2.4	Valores de datos y propiedades <i>datatype</i>	129
7.3	Recapitulación y reflexión	130
8	METODOLOGÍAS DE REPRESENTACIÓN DE ONTOLOGÍAS: METHONTOLOGY	132
8.1	Recapitulación y observaciones	139
9	HERRAMIENTAS PARA DESARROLLAR ONTOLOGÍAS	141
9.1	Protégé 4	141

9.1.1	Interfaz del programa	142
9.1.2	Creación de clases	146
9.1.3	Creación de propiedades	148
9.1.4	Creación de instancias	149
9.1.5	Descripción de clases	150
9.1.6	Descripción de instancias	152
9.1.7	Descripción de propiedades	154
9.1.8	Multilingüismo y polisemia	156
9.1.9	Creación de axiomas generales	156
9.1.10	Punning	158
9.1.11	Consultas	158
9.1.12	Razonadores para OWL 2 2 en Protégé	158
9.2	TopBraid Composer	160
9.3	Recapitulación y observaciones	163
10	CONCLUSIONES GENERALES Y SÍNTESIS DEL MARCO TEÓRICO DE NUESTRO TRABAJO EMPÍRICO	164
III.	DESCRIPCIÓN DEL TRABAJO EMPÍRICO	166
1	INTRODUCCIÓN AL TRABAJO EMPÍRICO	167
2	RECURSOS TEXTUALES	168
2.1	Normas UNE de cerámica	168
2.1.1	Norma UNE-EN 14411	168
2.1.2	Norma UNE-EN-ISO 10545	172
2.2	Catálogos comerciales de cerámica	176
2.3	El corpus TxtCeram	178
2.4	Descripción de la baldosa cerámica de PROALSO	178
2.5	El Diccionario Cerámico Científico-práctico	179
3	HERRAMIENTAS INFORMÁTICAS	180
3.1	WordSmith	180
3.2	Protégé 4 y HerMiT	181
4	IDENTIFICACIÓN, DESCRIPCIÓN Y ESTRUCTURACIÓN DE LOS ELEMENTOS QUE AYUDAN A DESCRIBIR UNA CARACTERÍSTICA	186
4.1	Introducción	186

4.2	Característica	188
4.2.1	Nombre de la característica	188
4.2.2	Valor de la característica	189
4.2.3	Propiedades matemáticas de las características	192
4.3	Ámbito de especialidad	192
4.4	Esencial, no esencial o Distintiva	193
4.5	Intrínseca o extrínseca	196
4.6	Relación superordinado–subordinado y subordinado–superordinado	197
4.7	Relación de meronimia	199
4.8	Relación característica- proceso	201
4.9	Características de las características	202
4.9.1	Característica equivalente	202
4.9.2	Instrumento de ensayo	202
4.9.3	Parámetro de ensayo	203
4.9.4	Unidad de medida	203
4.10	Sinonimia y equivalencia en otra lengua	203
4.11	Otra información en lenguaje natural	204
4.12	Propuesta de modelo de descripción de una característica	204
5	EXTRACCIÓN DE LAS CARACTERÍSTICAS DE LA BALDOSA CERÁMICA Y SU INFORMACIÓN CONCEPTUAL	209
5.1	Extracción de una lista inicial de candidatos a característica del producto acabado.	209
5.2	Extracción de información conceptual sobre las características.	210
5.3	Recopilación de descripciones sobre tipos de baldosas	213
5.3.1	Clasificación de la norma UNE-EN 14411	214
5.3.2	Clasificación comercial de las baldosas cerámicas de PROALSO	216
5.3.3	Resultados y observaciones	217
5.4	Consulta a experto y validación	217
5.5	Resultados	222
6	FORMALIZACIÓN DE LAS CARACTERÍSTICAS EN LA ONTOLOGÍA	224
6.1	Formalización con Protégé en OWL 2	224
6.2	Formalización de clases en la ontología y aspectos generales	224

6.3	Formalización de característica: nombre y valor	232
6.3.1	Características con valores de tipo unidad léxica	233
6.3.2	Características con valores booleanos	236
6.3.3	Características con valores numéricos sin rango limitado	237
6.3.4	Características con valores numéricos con rango limitado	237
6.3.5	Característica con una expresión matemática compleja como valor	238
6.3.6	Observaciones	239
6.4	Formalización de características equivalentes	240
6.5	Formalización de otra información en lenguaje natural	242
6.6	Formalización de las relaciones en la ontología	244
6.6.1	Relación baldosa cerámica – proceso de fabricación	244
6.6.2	Relación característica – proceso de ensayo	244
6.6.3	Relación de meronimia	245
6.7	Formalización del tipo de característica intrínseca o extrínseca	246
6.8	Formalización de características esenciales y distintivas en la ontología	248
6.8.1	Descripción de las clases de característica de la baldosa cerámica	249
6.8.2	Descripción de las clases de baldosa cerámica	252
6.9	Formalización de la sinonimia y equivalentes a otras lenguas	254
6.10	Resumen de los elementos que componen la ontología	256
7	INTRODUCCIÓN DE DATOS EN LA ONTOLOGÍA	261
7.1	Creación de clases y estructura de clases	261
7.2	Introducción de características en forma de <i>object property</i> e introducción de valores	264
7.2.1	Casos particulares	267
7.3	Introducción de características en forma de <i>datatype Property</i> e introducción de valores	268
	Casos particulares	272
7.4	Creación de características y relaciones para describir las características de la baldosa	273
7.5	Descripción de las clases de característica	273
7.5.1	Tipo de característica	274
7.5.2	Introducción de relaciones, características esenciales y características distintivas	276
7.5.3	Contextos	280
7.5.4	Fuente de la característica y fuentes complementarias	281
7.6	Descripción de las clases de baldosa	281
7.7	Descripción de otros elementos en la ontología	286
7.8	Introducción de equivalentes y sinónimos	287

7.9	Resultados de la introducción de datos en la ontología	288
8	LA ONTOLOGÍA DE CARACTERÍSTICAS DE LA BALDOSA CERÁMICA	289
8.1	Prueba de la inferencia y la consistencia en la formalización de características de la baldosa cerámica	289
8.2	Ejemplos de consultas en la ontología	291
8.3	Análisis de la clasificación inferida de baldosas cerámicas	293
IV.	GENERAL CONCLUSIONS	296
	BIBLIOGRAFÍA	303
	GLOSARIO DE ACRÓNIMOS Y ABREVIATURAS	313
	ÍNDICE DE FIGURAS	314
V.	ANEXOS	320
	ANEXO 1 - ACCESO A LA TABLA DE DESCRIPCIÓN GENERAL DE CARACTERÍSTICAS	321
	ANEXO 2 - FICHAS DESCRIPTIVAS DE BALDOSAS	322
	ANEXO 3 - LISTADO DE INSTRUMENTOS	326
	ANEXO 4 - LISTADO DE PARÁMETROS	329
	ANEXO 5 - LISTADO DE PROCESOS	336
	ANEXO 6 - LISTADO DE UNIDADES DE MEDIDA	340
	ANEXO 7- GLOSARIO DE EQUIVALENCIAS EN ESPAÑOL E INGLÉS DE LAS CARACTERÍSTICAS Y BALDOSAS	342
	ANEXO 8 – GRAFO DE LAS SUPERCLASES DE LA ONTOLOGÍA	345
	ANEXO 9 - GRAFO DE LA CLASE ‘CARACTERÍSTICA DE LA BALDOSA CERÁMICA’	347

ANEXO 10 - BIBLIOGRAFÍA DE RECURSOS TEXTUALES 349

**ANEXO 11 – INFORMACIÓN PARA LA INSTALACIÓN DEL SOFTWARE Y
ACCESO A LA ONTOLOGÍA 356**

I. Introduction

1 Overview

Everyday society manages a significant volume of information. On the one hand, it results of a need for interchanging this information. On the other hand, this is the result of an extraordinary development of means and resources that allow carrying on information interchange. In this context, our work is based in the idea that reusing information and creating knowledge bases will contribute to efficiently manage resources.

Research on language is sometimes oriented to tackle the needs of information society. Among this type of studies, we find those works addressed to improve information interchange among experts. In the context of data interchange, we believe that the study of the elements related to the use of terminology on a domain of a field of expertise should help to ensure success in communication.

Terminology and the use of an interdisciplinary approach where computing, documentation or linguistics are included, provide a necessary support to improve information interchange globally. Authors such as Temmerman (2000), Faber (Faber et al., 2005) or Roche (2005) contribute with their views to facing up the new needs of terminologists and the challenges demanded by modern terminology in the information society. The systematization of terminological resources or terminology-related automatic processes have become very important aspects to improve information interchange and working (Alcina, 2001).

Regarding the technology developed to meet these goals, research improvements comprise a lot of applied domains. Processes such as terminological data extraction for texts are changing from manual to automatic techniques, which minimize efforts (Feliu, 2004). Furthermore, knowledge bases and ontology systems have emerged as powerful concept tools able to manage terminological data and knowledge (Gruber, 1995; Guarino, 1995). In this regard, the possibility of using ontology reasoners takes our interest. Ontology reasoners are computing applications that analyse the information represented in the ontology and are able to infer new information that has not been represented explicitly in the ontology (Bock et al., 2008).

In general and in most of the cases, research projects found on this topic are based on inter-disciplinary approaches, among which we find terminology and knowledge engineering (Estellés, 2012). Nowadays, in the studies on ontologies and terminology, emphasis is on improving the tools used by experts of language and information (such as electronic encyclopaedias, terminological management systems or assisted translation applications).

We are aware of the need of researching in systems that improve access to terminology and ensure a fasten creation of the resources that enable communication among experts. This study aims to contribute on this topic with a research on representation and formalization of concept properties in the form of ontology. TxtCeram¹ and Ontodic² research projects are the background of this research.

¹ «TxtCeram: Extracción semiautomática y análisis conceptual formal de términos de la cerámica a partir de un corpus electrónico» (semi-machine extraction and formal conceptual analysis of tiling terminology based on electronic corpora). This research project is funded by the regional government of Valencia (code GV05/260) and directed by Amparo Alcina.

TxtCeram (Alcina, 2006) is a research project developed by the Tecnolettra team (Universitat Jaume I). The project focuses on the automation of extraction processes and on formal conceptual analysis of the terminology of ceramics. Main project goals were the test of the efficiency of computing tools used to design a system for assisted terminology, the study of their benefits for linguistic mediation and the test of ontology edition and its prospects for knowledge bases generation.

Ontodic (Alcina, 2008) is, to some extent, the continuation of TxtCeram Project. The main goal was to design a methodology for creating onomasiological dictionaries. Onomasiological dictionaries include terminology of a domain and allow users to do queries from the meaning of the term and not only from its name. This kind of dictionaries is addressed to linguistic mediators (translators, interpreter, technical writers...) though it may be interesting for experts and wider public as well.

Our work is focused on the formalization of concept properties in the form of ontology. From a terminological point of view, concepts can be classified in four types of categories: entities, activities, relationships and properties (Sager & Kageura, 1994: 193). Properties are a subject of study crucial to describe concepts since they help to distinguish among concepts. Properties can be of several types. Thus, we can find properties that refer to different aspects. For instance, in the domain of ceramics, we find properties that refer to the colour, the shape or the function of a tile. Furthermore, not all of them are equally important when defining a concept. Sometimes, properties serve to distinguish among concepts and sometimes properties serve to group concepts according to them. Concept description implies the study of the “range”, “importance”, “nature” or “origins” of properties.

We aim to focus on the type of information provided by a property. It involves the information required to describe properties when they are used to describe concepts and the information required to describe properties as a concept itself. We will study how does a property contribute to the formalization of other concepts and how can it be formalized and represented in other to provide information on itself. The raw material of this research is ceramics property, and more specifically, ceramic tile properties.

²«ONTODIC I: Metodología y tecnologías para la elaboración de diccionarios onomasiológicos basados en ontologías. Recursos terminológicos para la e-traducción» (methodologies and technologies for creating onomasiological dictionaries based in ontologies. Terminological resources for e-translation). This research Project is funded by the Department of Education and Science of Spain, (code TSI2006-01911) and « ONTODIC II: Metodología y técnicas para elaborar diccionarios de colocaciones basados en ontologías. Recursos terminológicos para la e-traducción» (Methodology and techniques for creating collocation dictionaries, Terminological resources for e-translation). This research Project is funded by the Department of Science and Innovation of Spain, (code TIN2009-07690, subprogram TSI) and directed by Amparo Alcina.

2 Assumptions and goals

The assumption on which this research is based is that properties of a domain of speciality, in this case of ‘ceramic tile’, can be described according to the criteria used to classify characteristics found in the bibliography of Terminology, Lexicography and Knowledge Engineering. This conceptual information can be formalized to (formally) represent properties in ontology.

Our main goal is representing the properties of a domain of speciality in a system for ontologies by using a standard ontology representation language and considering the criteria of classification of properties found in the bibliography on Terminology, Lexicography and Knowledge Engineering. The outcome of this research is targeted to be used as a searching tool for terminologists or language mediators interested in a way for representing concepts and in the representation of ceramic tile properties and their nature.

Specific targets are:

- Identification of domain tile properties and compilation of conceptual information by using a knowledge-based approach;
- Gathering a typology of ceramic tiles by using a text-based approach;
- Building a formalization model of properties;
- Designing a methodology for representing properties in an ontology editor;
- Testing conceptual inference applied to property representation;
- Comparing the knowledge-based approach used to represent properties and the text-based approach used to represent ceramic tiles.

3 Methodology

One of the main challenges of this research has been to find the balance between the terminological approach, the technical side involved in the development of an ontology and the study of the domain of ceramics; all of this avoiding technicalities and theorisations.

To address the description of ceramic tile properties we have analysed the criteria of classification and description existing in the bibliography of Terminology and Linguistics. Those criteria and contributions have been complemented with more general contributions on concept description found in the bibliography of Knowledge Engineering and other areas.

Then, we have extracted information about ceramic tile properties attending to the results of the previous analysis and to the field. In order to do that, we have extracted information for several text sources on ceramics and ceramic tile and we have been assessed by an expert. During this stage we also have extracted equivalents in English of some of the terms, mainly the properties, and a typology of tiles using a text-based approach.

First and second stages were carried out simultaneously. The analysis of the bibliography helped us to identify the elements that might be needed to describe a property. At the same time, during the extraction and data collection stage on tile properties, the definition of the elements used in the description acquired a new complexity that has not been detected in the analysis stage.

The approach used in the case of the typology of ceramic tiles is different to the one used in ceramic tile properties. A classification of types of ceramic tiles has been extracted from what we found explicitly in the bibliography, without adding any detail, contrasting with experts or inferring any underlying knowledge. As a result we have created a series of datasheets that include the elements that take part in the description of a property. Those datasheets and work sheets are based in the aspects identified during the analysis and, in the case of ceramic tile properties, also in the information learned when extracting conceptual information about them. It is a compilation of all the aspects to be considered when creating an ontology and it includes the relationships that among the elements. At the end of these stages the following goals were accomplished:

- Identifying ceramic tile properties and compiling their conceptual information using a knowledge based approach;
- Collecting a typology of ceramic tiles using a text-based approach.

Designing a formalization model for properties has been the following aim. The methodology used to do so is based in the analysis, from a technical point of view, of the information we wanted to represent in the ontology, data types and relationships among data types. And thus, proposing a valid formalization schema.

Our aim has been to be fair to the information we wanted to represent and finding the balance between our goal and the technical prospects of the technology used. This stage has been reviewed several times during the testing of data implementation with the help of an ontology reasoner.

This work ends with the implementation of the data gathered during the extraction stage. Implementation of data has followed the formalization system designed in the previous stage. Data has been implemented in the ontology manually and using automatic insertion techniques. The formalization model and data implantation served us to set up and refine a methodology for representing properties in an ontology editor.

Finally, a reasoner has been used to check ontology consistence and to test some basic queries that can be answered with the use of the ontology and logics. Furthermore, we have compared the consistence of properties classification and ceramic tiles classification. The goal was to verify if the ontology was consistent enough, check for any lack of relevant information in the field and detect it. The analysis and results have been developed and some conclusions have been made regarding the methodology and theoretical aspects. The result proves the benefits of using conceptual inference in the representation of properties and allowed us to compare, according to the inferred data, the approach used to represent properties against the approach used to represent ceramic tiles.

4 Structure

The present study is divided in four main parts. The first part introduces the research done. The second part presents the theoretical foundation used as a basis for developing this study. The third part presents and describes the empirical study and finally, the fourth part presents the conclusions of this research.

Focusing on the second, third and fourth part. Firstly (§II), we analyse different theoretical approaches that help as a basis for the empirical study. We focus, mainly, in the bibliography on Terminology but we also approach to Linguistics and Knowledge Engineering with the prospects of finding new points of view and technologies.

Chapter 1 of this second part of theoretical foundations begins with the analysis of different theories of Terminology that inspired us to develop an ontology on ceramic tile properties. Next, in chapter 2 (§II.2), we focus in the concept of property and we analyse its meaning with the help of different authors. We present some other concepts needed to understand the meaning of property, such as *term* or *intension*, to then go to properties, types of properties and relationship between properties and other concepts. In chapter 3 (§II.3), we study the role of properties in concept representation, we explore the prospects for representing concepts. Chapter 4 (§II.4) presents some of the methodologies available in the bibliography on concept representation in Terminology. Finally, chapter 5 (§II.5) describes some tools used in some projects to create ontologies.

Regarding Knowledge Engineering, chapter 6 (§II.6) begins a tour that starts with a short introduction to the key concepts of this area and leads us to focus in ontologies, where we present some examples of ontologies. In chapter 7 (§II.7), we explore the basics of Ontology Web Language 2 (OWL 2) and its prospects. Chapter 8 (§II.8) presents a methodology for building an ontology and finally, chapter 9 (§II.9) focuses in the editing tools and other available resources available in Knowledge Engineering to create ontologies.

The third part, empirical research description, begins with an introduction to the empirical research (§III.1), a description of the textual resources (§III.2) and of the tools used to carry out the study (§III.3). Chapter 4 (§III.4) describes the analysis of the aspects used to describe and classify concept properties, the analysis is based in different studies on Terminology, Linguistics and Knowledge Engineering areas. At the end of this chapter 4, the results of the analysis are presented, consisting in a list and description of the elements that we have used to describe and classify properties.

Then, in chapter 5 (§III.5), we describe the process for extracting ceramic tile properties and their conceptual information. We also show how we gathered this information considering the description, classification and validation criteria. Furthermore, we describe the process used to gather a typology of tiles and their differences regarding properties extraction.

In the next chapter (§III.6), we address the design of a model for formalizing properties in the ontology. The formalization model includes the elements used to describe a property, its specifications and prospects. It is formalized in OWL 2 with the help of Protégé OWL 4.1 editor. Equally, we analyse the different ways to represent each element and its relationships, the developments detected while applying OWL 2 syntax and present a proposal of formalization according to the needs of this study.

Chapter 7 (§III.7) presents the methodology used to implement all the data in the ontology, some problems found during the implementation and some solutions. Finally, chapter 8, analyses the consistency of the formalization of ceramic tile properties and the formalization of tiles. This chapter also includes some examples of queries that can be done with the help of an ontology reasoner. General conclusions are presented in the fourth part.

II. Fundamentos teóricos

1 Enfoques en el estudio de la terminología

Este capítulo explora distintos acercamientos a la terminología. Partimos de la Teoría General de la Terminología para centrarnos en las aportaciones que abogan por un enfoque más interdisciplinar de la terminología. Situadas en este marco, nos adentramos en cuatro enfoques orientados a la representación de conceptos siguiendo un orden cronológico: el trabajo de Meyer y colegas, la Terminología Socio-cognitiva, la Terminología Basada en Marcos y la Ontoterminología.

1.1 La Teoría General de la Terminología vs la interdisciplinariedad de la terminología

La terminología se ocupa del estudio y la recopilación de términos especializados (Cabré, 1992: 17). Alrededor de los años 30 empieza a tomar forma como disciplina gracias al trabajo de Wüster y la Escuela de Viena (Wüster, 1998) que culmina en lo que se ha venido a llamar la Teoría General de la Terminología (TGT). Esta primera teoría en la historia de la Terminología trata principalmente de la sistematización de los métodos de trabajo en terminología, los principios que deben considerarse en el estudio de los términos y una metodología para tratar estos datos terminológicos.

Para ello, en la TGT se establecen un conjunto de principios que pretenden dotar a la disciplina de una capacidad normalizadora y organizadora de la que hasta entonces carecía. Por eso, se hablará de que la terminología debe centrarse en el estudio de los términos y de que éstos pertenecen a un lenguaje específico de un dominio concreto. El foco de atención se fijará en el concepto y no en el término, considerado algo independiente del concepto. La TGT entiende el término como una ‘denominación’ de un concepto y el concepto como un elemento del pensamiento, una idea. Entre los principios de que deben considerarse para estudiar los términos según la TGT, destacan:

La aplicación de una *perspectiva onomasiológica* en la elaboración de las terminologías. La TGT se centrará en el concepto y no en las palabras. Su función será explicar los conceptos de un dominio de especialidad, asumiendo que puede haber casos en los que dichos conceptos no puedan identificarse con una palabra. Esto contrasta con la perspectiva semasiológica que hasta entonces predominaba por influencia de la lexicología, donde los diccionarios se centran en la descripción de las palabras admitiendo que una palabra puede tener más de un significado.

El concepto se observa como el eje de estudio y como algo independiente de la denominación. Esta idea se basa en la perspectiva onomasiológica que acabamos de mencionar y defiende la naturaleza independiente del concepto (independiente del lenguaje).

El principio de univocidad. La TGT asume que cada concepto estará designado por un único término y un término designará a un único concepto. Este principio nace de la preocupación por la estandarización en terminología y recomienda evitar en la manera de lo posible caer en la normalización de más de una término (denominación) para un concepto.

La naturaleza prescriptiva de la Terminología. Para cumplir su propósito de univocidad y con el fin de ahorrar esfuerzos, la TGT recomienda que la Terminología centre su labor en la normalización terminológica.

La importancia de las características de los conceptos. En la TGT se establece que todo lo que al analizar un concepto se percibe como el conjunto de sus características individuales se denomina *intensión conceptual*. Es decir, la suma de características individuales de un concepto será el mismo concepto.

Esta teoría ha sido criticada posteriormente por numerosos autores para adaptarla y dar respuesta a las nuevas necesidades que ha afrontado la disciplina. Uno de los aspectos más criticados es que la terminología no solo se debe ocupar de la normalización y prescripción sino también de la descripción (Wright & Budin 1997: 329). La Terminología está presente cuando se trata de documentar y recopilar los términos de uso de una determinada comunidad especializada con fines de formación o para mejorar la comunicación entre los especialistas y en este aspecto, se requiere de la descripción de los términos y no tanto de la normalización de éstos. Si bien es cierto que existen organismos que se ocupan de la normalización, las nuevas corrientes teóricas señalan que el trabajo de normalización terminológica siempre debe ser posterior al trabajo descriptivo de manera que se garantice el uso de la terminología normalizada. Esto evidencia la dificultad práctica para llevar a cabo el principio de la univocidad. No siempre se dará el caso de que exista una única denominación para un concepto si tenemos en cuenta que es la comunidad especializada la que influye en las normas de uso de un término y no siempre se podrá atender a aspectos prescriptivos o normalizadores.

Con el tiempo, estas y otras aportaciones críticas han surgido de la mano de distintos autores al notar que la TGT no atiende a las necesidades reales de la comunicación especializada. Cabré (Cabré, 2002) considera que esta situación se explica porque la Terminología como disciplina interdisciplinar también se compone de aspectos lingüísticos, cognitivos y sociales que la TGT no había considerado.

Autores como Sager (Sager, 1993) o Cabré (ibíd.) contribuyen a que la disciplina de la Terminología se constituya plenamente en toda su faceta interdisciplinar.

Sager (ibíd.: 12) señala que la terminología, al estar basada en aspectos semánticos, podrá abordarse desde tres puntos o tres dimensiones: la lingüística, la cognitiva y la comunicativa. Las tres dimensiones que identifica Sager en la teoría de la Terminología se podrían explicar brevemente a partir de los siguientes puntos:

- La dimensión cognitiva será la que relaciona las formas lingüísticas con el contenido conceptual (el referente en el mundo real)
- La dimensión lingüística examina las formas de representación potenciales y existentes de terminologías (la designación)
- La dimensión comunicativa observa el uso de las terminologías y justifica la actividad humana de compilación y procesamiento de terminología (el uso de referente y designación)

Según esto, el estudio terminológico podrá abordarse también desde tres puntos de vista (Sager, 1993: 13):

- El referente en el mundo real

- La designación dada al referente
- El uso de terminologías

El concepto dejaría de ser el foco exclusivo de estudio y compartiría la atención con el término y la situación comunicativa.

El enfoque de Cabré (Cabré, 2002) coincide y profundiza en el planteamiento de Sager. Esta autora, en su revisión de la TGT, aportará algunas observaciones al respecto a través de su teoría comunicativa de la terminología (TCT). La teoría señala que los términos son unidades formadas por tres vertientes diferentes: la vertiente semiótica y lingüística, vertiente cognitiva y vertiente comunicativa.

De acuerdo con esto, se puede acceder al estudio terminológico también desde diferentes vertientes. Por ejemplo, podemos partir del conocimiento para llegar a la materia lingüística y a la comunicativa o de la materia lingüística y la comunicativa para acceder al conocimiento.

De este modo, aspectos como la variación de denominaciones para un mismo concepto se explicarán desde la vertiente lingüística como un fenómeno lingüístico que el terminólogo deberá considerar en su trabajo. Según Cabré, la comunicación especializada se caracteriza principalmente por factores de tipo lingüístico (semánticos, léxicos y textuales) y de tipo pragmático (emisor, destinatario...). Además, dependiendo de estos factores, la comunicación especializada permite establecer diferentes niveles de especialización y diversos grados de lo que se denomina *opacidad cognitiva*, es decir, en la comunicación especializada encontraremos diferentes índices de densidad terminológica y cognitiva (Cabré, *ibíd.*).

1.2 Representación de conocimiento en Terminología según Meyer

El trabajo de Meyer (Meyer et al., 1992a; Meyer et al., 1992b; Meyer y Skuce, 1997) se considera pionero entre las nuevas corrientes orientadas al concepto. A través del proyecto Cogniterm, Meyer y otros proponen representar el conocimiento de un ámbito a través de bases de conocimiento terminológico. Esta propuesta se justifica bajo la vertiente conceptual de la terminología, en la que la investigación se centra en desarrollar herramientas que ayuden a facilitar el trabajo de representación de conceptos en Terminología. En este contexto, Meyer es la propulsora de las bases de conocimiento terminológico concebidas como un nuevo método para representar las relaciones que se establecen entre los conceptos de un ámbito de especialidad (véase §3.4.1.1).

Meyer se centra en la dimensión conceptual e identifica los momentos en los que esta dimensión puede resultar más relevante a lo largo del trabajo terminológico. Entiende *dimensión conceptual* como la dimensión vinculada a adquirir, analizar, comparar, formalizar y revisar (gestionar) información sobre los conceptos de los términos. En su trabajo, persigue identificar los aspectos presentes en las actividades diarias del terminólogo vinculados a la dimensión conceptual. Y una vez identificados, automatizar parcial o completamente estas actividades a través de herramientas específicas como las bases de conocimiento terminológico.

La autora se basa en cuatro actividades generales que forman parte del trabajo del terminólogo con el fin de describir, a partir de estas, las sub-actividades vinculadas a la dimensión conceptual de la terminología (Meyer, 1993: 140-151):

- selección de la documentación

- delimitación de la terminología
- elaboración de entradas terminológicas
- control de la calidad

Estas categorías representan procesos o etapas que tienen lugar durante la realización de un trabajo terminológico bilingüe sobre un dominio concreto y donde el terminólogo no siempre es experto en la materia a tratar.

La primera de las actividades generales que menciona Meyer es la de **selección de la documentación**. Antes de analizar o seleccionar los términos, los terminólogos deben partir de un corpus documental que servirá como fuente de conocimiento del proyecto. Otra ayuda pueden ser los expertos del ámbito u otros terminólogos expertos.

Desde el punto de vista conceptual, el proceso de familiarización con el ámbito a trabajar se lleva a cabo en dos niveles que Meyer denomina *macro-nivel* y *micro-nivel*. En el macro-nivel se atiende a las estructuras del conocimiento general y en el micro-nivel a las características conceptuales de conceptos individuales.

En el macro-nivel, el terminólogo empieza a familiarizarse con el conocimiento general del ámbito, intentando establecer sus límites, divisiones, áreas que se solapan con otros campos etc. En esta fase, el terminólogo puede realizar esquemas sobre el conocimiento trabajado mediante redes conceptuales, por ejemplo. En el micro-nivel, los terminólogos, se centran en aquellos conceptos individuales que destacan en los textos por su relevancia con el ámbito.

Las aplicaciones prácticas de estas actividades ayudan a clarificar el ámbito del proyecto. Al tener una idea clara de cuáles son los límites conceptuales del campo podemos delimitar el rango de documentación que necesitamos. Al determinar las áreas que se solapan con otros campos, podemos establecer enlaces a documentación relacionada. La familiarización con las micro-estructuras y las macro-estructuras del campo permite que los terminólogos conozcan los puntos de entrada a la documentación ya que los nombres de los sub-campos, conceptos clave y características de estos conceptos puede servir como palabras clave de la búsqueda.

El hecho de tener una idea general de la estructura jerárquica del campo de especialidad también ayuda al proceso de selección de la documentación en lo relativo a ordenar la bibliografía de más general a más específica. Otro aspecto positivo es que ayudará a organizar la documentación a utilizar, permitiendo distinguir entre los distintos sub-ámbitos para cada lengua.

Finalmente, estas actividades de tipo conceptual proporcionan a los terminólogos un marco conceptual y la terminología básica necesaria para realizar una comunicación efectiva con los agentes que proporcionan la documentación y con los expertos del ámbito. En lo que se refiere a los expertos, también ayuda a la comunicación eficiente y directa.

La **delimitación de la terminología** es la segunda actividad que estudia Meyer y consiste en identificar los términos potenciales y sus contextos, así como resolver los problemas que ello conlleva (contextos inadecuados, ausencia de términos para conceptos reconocidos...). Este proceso incluye agrupar los datos sobre cada término (sinónimos, etc.)

Las actividades conceptuales en esta fase tienen que ver con el desarrollo de los esquemas conceptuales que habían elaborado en el punto anterior. También se inicia otro proceso de tipo

cognitivo: el análisis de las características de los términos individuales basándose en los contextos encontrados.

Las aplicaciones prácticas de estas actividades conceptuales implican que el terminólogo será capaz de identificar elementos léxicos que constituyen términos. Al entender las relaciones conceptuales podrá identificar, por ejemplo, términos complejos.

El proceso de identificación de términos, incluye eliminar aquellos que se consideran *ruido* (falsos términos, términos de otros ámbitos, etc.). Asimismo, los terminólogos buscarán posibles lagunas sobre sub-campos de los que se carece de términos o que no se habían tenido en cuenta previamente.

Llegado el momento de organizar los datos (términos y contextos), el razonamiento sobre las características conceptuales de éstos ayudará al terminólogo a evaluar la calidad de los contextos y a determinar cuáles deben eliminarse o substituirse o a identificar posibles casos de sinonimia que ayuden a la reagrupación de términos. Mientras el terminólogo se prepara para determinar qué términos formaran parte del trabajo, alimenta su conocimiento sobre el marco conceptual del ámbito lo que puede mejorar la comunicación con el experto y el documentalista.

La tercera actividad alude a la **preparación de entradas terminológicas** para los términos identificados en el punto anterior, que se describe como el análisis sistemático de los términos en su contexto y a la redacción de las entradas terminológicas. La principal función de este análisis es determinar los significados de los términos, aunque también sirve para identificar otras características de tipo lingüístico como el género, la frecuencia, etc.

Las actividades conceptuales en este punto se centran en la adquisición de un conocimiento más profundo que resulta necesario para desarrollar las entradas terminológicas. En un macro-nivel, esto implica establecer las separaciones conceptuales más relevantes en el ámbito y también determinar las relaciones jerárquicas y de otros tipos. En un micro-nivel, el análisis contextual implica la identificación de un número de características para cada contexto. Estas características se compararán con las de otros conceptos relacionados (posibles sinónimos, equivalentes en otra lengua) para poder determinar si son necesarias para establecer correspondencias entre términos.

Las aplicaciones prácticas de estas actividades resultan fundamentales, según Meyer, ya que el entendimiento de las estructuras conceptuales de un ámbito es crucial para determinar aspectos con el campo de dominio del término, los casos en los que un mismo término aparece en distintos ámbitos (de modo que prevenimos la repetición de entradas). Hay que tener en cuenta que cuando se trabaja con distintos campos de especialidad, los terminólogos suelen indicar el más básico, lo que implica tener un conocimiento profundo de las estructuras conceptuales involucradas.

En la construcción de definiciones, Meyer considera que es fundamental que los terminólogos entiendan las características del concepto que describen. En caso de elaborar una definición intensional clásica (*genus-differentia*), los terminólogos necesitan comparar estas características conceptuales con otros conceptos del mismo nivel en la jerarquía (conceptos coordinados) para identificar las características distintivas (*differentia*). Además de las relaciones jerárquicas, deben analizar otro tipo de relaciones (*parte-todo*, etc.) y reflejarlas en la definición. Cuando se usa una definición de tipo contextual, el contexto que más características y relaciones proporciona será el más adecuado.

Finalmente, en lo que refiere al **control de la calidad** donde se incluye la revisión y actualización del trabajo, mientras se elabora y una vez finalizado.

Las actividades conceptuales en este punto se centran en aspectos de análisis y debate sobre problemas conceptuales identificados por expertos o terminólogos. Asimismo la actualización periódica del trabajo, implica un control de cambios en las estructuras de conocimiento y las características conceptuales.

Las aplicaciones prácticas de estas actividades están vinculadas al entendimiento de ámbito, algo necesario para poder interpretar las observaciones de los expertos y utilizar la información sobre estas observaciones. En las actualizaciones, es necesario entender profundamente las estructuras de conocimiento existentes para poder modificarlas con características conceptuales o estructuras nuevas.

Aún así, Meyer, asume que los problemas conceptuales son sustancialmente más complejos que los puramente lingüísticos y la tecnología necesaria para enfrentarse a ellos todavía es muy limitada. No obstante, ya en 1993 señalaba que los avances en el campo de la inteligencia artificial hacen que la gestión de conceptos en terminología con la ayuda de herramientas informáticas empiece a ser algo cada vez más posible (Meyer, 1993: 151).

De este enfoque que hace Meyer del trabajo del terminólogo visto desde una dimensión conceptual destacamos las continuas aplicaciones prácticas que hace la autora de las observaciones teóricas. En el trabajo de Meyer, todas las reflexiones teóricas tienen una relación directa con la realidad práctica y pueden servir de guía para planificar un trabajo terminológico atendiendo a la dimensión conceptual de la terminología.

1.3 Terminología Socio-Cognitiva

La Teoría Socio-cognitiva de la Terminología (Temmerman, 2000) nace con el propósito de ofrecer un marco teórico y de trabajo alternativos a los que se desprenden de la Terminología tradicional (Wüster, 1998). Su propuesta considera que los principios de la Terminología tradicional se concentran en la estandarización y dejan de lado la descripción realista de los términos y su significado.

Temmerman propone un marco teórico con el que más tarde fundamentará su metodología, basado en la semántica cognitiva y la hermenéutica (Temmerman, 2000: 54-66) y. La Teoría Socio-cognitiva de la Terminología que se basa en cinco principios que se deben entender como postulados básicos de la teoría y que la autora distingue de los métodos de trabajo que agruparían herramientas o técnicas utilizadas en el trabajo terminológico (ibídem: 222-227). Temmerman explica estos postulados enfrentándolos a la Teoría General de la Terminología y a otros enfoques relacionados.

El *primer postulado* trata el significado de 'concepto'. Temmerman señala que la definición de 'concepto' de la norma ISO 1087 (ISO, 1995) es demasiado restrictiva. En esta norma el concepto se define como «unit of thought constituted through abstraction on the basis of properties of a set of one or more objects».

Temmerman argumenta que los conceptos no son algo objetivo que podamos identificar todos de la misma manera. En su opinión cada uno entendemos el mundo a través de modelos cognitivos o marcos, que en semántica cognitiva se conocen como modelos cognitivos idealizados (Lakoff, 1987: 285) y que se componen de unidades de conocimiento. Estas

unidades de conocimiento tendrán una estructura prototípica y están vinculadas entre ellas constituyendo los marcos. En ese sentido, Temmerman propone hablar de categorías cuando tratemos con unidades de conocimiento con una estructura prototípica que podamos identificar.

Al hilo de este argumento, Temmerman se plantea cómo podemos estudiar y describir una categoría o unidad de comprensión y propone hacerlo partiendo de los términos, ya que es lo que podremos observar a través del discurso y a través de ellos intentar discernir el conjunto de conocimiento al que se refiere, teniendo en cuenta que éste se puede reformular o modificar a lo largo del proceso de estudio.

El *segundo postulado* se centra en la estructuración de los conceptos. Considera que la comprensión de un concepto se basa en la categorización de unidades de comprensión y que éstas unidades tendrán estructuras intra-categoriales e inter-categoriales que hacen que la categorización pueda variar. En la TST, una categoría o unidad de comprensión no se considera independiente de la lengua. Los aspectos intra-categoriales aluden a las relaciones lógicas que se establecen entre categorías y a la relevancia al considerar los bloques de información, por ejemplo, propiedades, información histórica o procedimental. Estos aspectos variarán dependiendo del tipo de categoría de la que se trate: una entidad, una relación o categorías paraguas, las categorías que se utilizan para agrupar a otras. Desde una perspectiva inter-categorial, donde se consideran las relaciones de tipo ontológico de una categoría, la categorización dependerá de la perspectiva, el dominio y la intención del modelo cognitivo que se plantee. En otras palabras, existen aspectos inter-categoriales e intra-categoriales que influyen en cómo se organizan las categorías en un modelo u otro pero también en la comprensión de las categorías en sí mismas.

El *tercer postulado* aboga por una plantilla de comprensión que se compone de distintos módulos de información sobre la categoría a describir. Dependiendo del tipo de unidad de comprensión, se requerirá una información como esencial u otra.

El *cuarto postulado* defiende la funcionalidad y explica el uso de la sinonimia y la polisemia como parte de las necesidades funcionales de comunicación. Y finalmente, el *quinto postulado*, habla de la constante transición de los modelos cognitivos, que justifica debido a: la necesidad de comprender más y mejor una realidad, la interacción entre distintos usuarios de distintas lenguas y la evolución de las estructuras prototípicas de las unidades de comprensión.

Los principales puntos de divergencia de esta teoría con respecto a la Terminología tradicional (TGT) se resumen en lo siguiente (Temmerman 2000: 228), en la Terminología tradicional:

- El concepto es el punto de partida para el estudio terminológico, mientras que en la TST es el término.
- El concepto es algo que existe de forma objetiva mientras que en la TST se habla de unidades de comprensión que pueden ser percibidas en la realidad o no, pero siempre pueden ser concebidas en la mente.
- Los conceptos se estructuran en sistemas de conceptos basándose en relaciones lógicas u ontológicas, mientras que en la TST las unidades de comprensión se organizan dentro de *modelos cognitivos idealizados* (Lakoff, 1987: 285).
- La definición de un concepto refleja su posición en un sistema de conceptos organizado mediante relaciones lógicas u ontológicas, en la TST la definición

dependerá del tipo de unidad de comprensión e incluirá una información u otra según la unidad.

- Se cree en la univocidad y por tanto, en la monosemia y mononimia, en la TST se acepta la polisemia y la sinonimia como algo funcional.
- El lenguaje figurado (uso de metáforas, etc.) debe ser limitado, mientras que en la TST se considera como parte de la descripción terminológica.

Sobre la representación de conceptos, Temmerman propone combinar teorías y métodos de análisis terminológico multilingüe de la terminología socio-cognitiva con métodos y directrices de la ingeniería ontológica. Se distingue entre modelado conceptual a un nivel independiente de la cultura y el análisis específico de unidades de conocimiento dependiente de la cultura, y por tanto, considera las estructuras prototípicas de conocimiento.

La combinación de terminología, entendida como el uso y estudio de términos, y ontología, entendida como el estudio de categorías de cosas que existen o pueden existir en un dominio, da lugar a un enfoque que se centra en compilar conjuntos de vocabulario de lenguajes especializados con el objetivo de:

- Describir ontologías con información terminológica (multilingüe), donde *ontología* se define como «a knowledge repository in which categories (terms) are defined as well as relationships between these categories» (Temmerman y Koen, 2003: 3).
- Estructurar los recursos terminológicos haciendo uso de las ontologías.

Se explica que ambos campos tienen muchas cosas en común ya que, por ejemplo, al construir una ontología o compilar una base de datos terminológica, tanto los expertos en ontologías como en terminología parten de la identificación de objetivos, la restricción del dominio, la especificación de los requisitos del usuario y la adquisición del conocimiento del dominio para extraer y entender las categorías y los términos.

Para explicar este enfoque, la autora alude a distintos cambios que han repercutido en la disciplina de la terminología y que principalmente son:

- La gestión terminológica computacional. Los métodos de trabajo convencionales en terminología han cambiado a raíz de la existencia de corpus electrónicos y del software para la gestión terminológica.
- Un cambio en el enfoque lingüístico. La disciplina de la terminología ha virado de un enfoque tradicional centrado en el concepto y la estandarización a un enfoque orientado a la comunicación y centrado en el discurso.
- El cambio en las ontologías. Se ha pasado de las bases de datos terminológicas a las bases de conocimiento terminológico, rebautizadas de acuerdo con Temmerman como *ontologías*.

Fundamentalmente, Temmerman propone utilizar este enfoque como una metodología para la gestión de conocimiento y la representación de dominios específicos, que combina el dominio de los expertos con la información proporcionada en lenguaje natural y se ocupa, sobre todo, de afrontar problemas debidos a objetivos multilingües (véase §4.2.4).

Esta teoría ofrece una base para desarrollar trabajos de terminología donde el enfoque cognitivo se combina con aspectos comunicativos y discursivos. El enfoque permite acercarse al

concepto desde el término y reflexiona sobre cómo podemos estudiar y describir una categoría o unidad de comprensión teniendo en cuenta que el conocimiento al que refiere puede cambiar.

De esta forma, la descripción de conceptos se observa como algo dinámico y flexible. Y en este sentido, destacamos la afirmación de que una unidad de comprensión se puede describir de más de una forma según su naturaleza, la perspectiva tomada, el dominio...

1.4 Terminología Basada en Marcos

La Terminología Basada en Marcos (TBM) es una propuesta del grupo Lexicon que se centra en la organización de los conceptos y su aplicación al trabajo terminológico (Faber et al., 2005; Faber et al., 2006; Faber et al., 2007; Faber, 2009). Los principales focos de estudio de la Terminología Basada en Marcos son:

- la organización conceptual de un dominio
- proponer un enfoque que permita describir la naturaleza multidimensional de las unidades terminológicas
- la extracción de la información sintáctica y semántica a través del uso de corpus multilingüe

Parten de la premisa de que la conceptualización de un dominio especializado está orientada según el objetivo y depende, en cierto grado, de la tarea que se vaya a desarrollar. Para llevar a cabo el trabajo de modelado del dominio proponen utilizar los textos y otros recursos de conocimiento especializado y así obtener la información necesaria para reflejar las diferentes conceptualizaciones y constructos en distintas lenguas. También, proponen analizar la estructura del lenguaje con el fin obtener un inventario de relaciones conceptuales que estructuren estos conceptos.

Para estructurar los conceptos de un dominio se basan en la Semántica de Marcos (Fillmore, 1982; Fillmore y Atkins, 1992) pero orientando el concepto de marco a las necesidades de la Terminología. La finalidad es proporcionar un enfoque que permita al terminólogo mejorar la forma de entender y trabajar los textos de un dominio a través del análisis de los conceptos del dominio y de las relaciones conceptuales que se establecen. Se trata de crear representaciones que no dependan del lenguaje aunque para ello se parte de los textos de especialidad en las distintas lenguas, de los que se irán extrayendo los significados.

Para la TBM, los marcos constituyen un tipo de elemento de estructuración cognitiva basado en la experiencia que proporciona el conocimiento del entorno y justifica la necesidad para la existencia de palabras en un lenguaje así como las palabras que se usan en un discurso. Los marcos tienen la ventaja de hacer explícito tanto el potencial semántico como el comportamiento sintáctico de las unidades de lenguaje especializado. Las redes conceptuales se basan en la información que subyace en un evento del dominio, lo que genera patrones para las acciones y procesos que tienen lugar en un campo de especialidad así como las entidades que participan en este.

La semántica de marcos y su aplicación práctica considera que para entender el significado de las palabras, uno debe primero conocer el significado de las estructuras o marcos conceptuales que subyace bajo su uso. Esto, según señala Faber, se podrá aplicar también a las unidades del lenguaje especializado (Faber, 2009: 120).

Esta teoría también se acerca a otros enfoques como la teoría Socio-cognitiva e Temmerman y la Teoría comunicativa de Cabré. Como aspectos en común, encontramos el estudio de la naturaleza multidimensional de las unidades terminológicas así como el uso de corpus textuales multilingües para extraer la información semántica y sintáctica relevante para la descripción del ámbito de interés.

En el caso de la naturaleza multidimensional, su representación se articula a través de la idea de ‘dominio’, un concepto que en la TBM se somete a revisión desde la Terminología y la Lingüística. De acuerdo con Faber, se pueden observar dos tipos de acercamiento en este asunto, los cuales no son incompatibles: por una parte, el dominio hace referencia al área de conocimiento en sí mismo y por otro, a las categorías de los conceptos dentro del campo de especialidad.

En lingüística, el concepto de dominio es también algo ambiguo según la autora. Sin embargo, debido a la relación entre lengua y pensamiento, Faber considera que tiene sentido pensar que para entender mejor el lenguaje, es necesario entender los conceptos que designan las formas lingüísticas. En este sentido, Faber considera que los enfoques cognitivos suponen la ventaja de observar la estructura conceptual como una parte fundamental del lenguaje y sirven para enfrentarse a ese aspecto.

Sobre la estructuración de conceptos, Faber (ibídem: 122) señala que el modo en cómo debemos estructurar las categorías de un dominio es un tema recurrente y que existen distintas propuestas en lo que se refiere a su organización. Estas propuestas coinciden en afirmar que una categoría puede ser cualquier cosa y que existen pocas limitaciones en los modos de estructurar categorías y sus miembros.

Faber señala que la noción de dominio desde la lingüística cognitiva es algo relevante aunque indeterminado y se adhiere al enfoque de Goldberg (Goldberg, 1998: 205) que entiende que el mundo está dividido en tipos de eventos que se podrían clasificar. Según la autora estos tipos de eventos se corresponden con los arquetipos conceptuales que describe Langacker (Langacker, 1991: 294-298). En el caso de la organización de los tipos de evento, se basa en la semántica ya que entiende que la sintaxis resulta insuficiente para establecer diferencias en significados. Un aspecto que la autora critica en cuanto a este tipo de enfoques es la ausencia de descripciones exhaustivas sobre la estructura interna de las categorías de un dominio, algo que considera esencial ya que afecta y limita lo que denomina *mapeado entre dominios* y que ella tiene en cuenta su enfoque. Basándose en la Teoría de Prototipos (Rosch, 1978), Faber entiende que los conceptos se pueden representar dentro de un mapa formado por círculos concéntricos y que estos se deben situar más cerca o más lejos del centro según su vinculación o tipicidad con respecto a un dominio.

Como puntos más relevantes de esta teoría encontramos la representación de conceptos a través de lo que denomina como *eventos* de un dominio y en forma de marcos de modo que se permita la interacción entre los conceptos que comparten distintos eventos. Los términos estarán vinculados a los conceptos de dichos marcos permitiendo el uso de términos en más de una lengua. La base para elaborar los marcos será el objetivo del trabajo y los conceptos se estructuraran de acuerdo con relaciones ontológicas y lógicas. Además, los conceptos o categorías de cada marco estarán vinculados a definiciones terminológicas y a contextos de uso, e incluso, a imágenes que permitan acceder mejor al significado de los términos representados

(Prieto y López, 2009). Más adelante, en §3.1.4.3 y §3.1.4.4 mostramos aplicaciones a trabajos terminológicos de esta teoría.

1.5 Ontoterminología

La Ontoterminología es una propuesta del equipo Condillac que sitúa las ontologías en el centro de estudio de la Terminología (Roche, 2008: 14). Roche defiende el doble papel que supone este enfoque: para la construcción de sistemas conceptuales y para lo que denomina *operatividad (opérationnalisation)* de la terminología, que podría entenderse como la funcionalidad de la terminología. La ontoterminología insiste en la importancia de los principios epistemológicos propios que residen en la conceptualización de un dominio y al mismo tiempo, en la necesidad de un enfoque científico a la terminología donde la ingeniería juega un rol fundamental. Combina el concepto de ontología de la filosofía antigua al hablar de principios epistemológicos de una conceptualización y el concepto de ontología de la Ingeniería del Conocimiento. Uno de los puntos más relevantes de esta teoría es que distingue entre la definición de un término y la definición de un concepto y lo justifica en base a las diferencias entre la dimensión lingüística y la dimensión conceptual de la Terminología (Roche et al., 2009) que dan lugar a distintos tipos de lenguaje (la lengua de especialidad, la lengua de la intelección y los lenguajes de representación).

Roche (Roche, 2008: 8) identifica tres tipos de lenguaje que tienen cabida en la práctica terminológica:

- La lengua de especialidad (lengua de uso).
- La lengua de la intelección.
- Los lenguajes de representación.

Respecto a la lengua de especialidad, señala que la Terminología a través de la Lingüística estudia los términos que se usan en un discurso, la relación entre los términos de uso y los significados en función de un determinado contexto. Centrarse únicamente en el discurso supone olvidar que ante todo la Terminología es una actividad científica que consiste en comprender, modelar y representar una realidad y razonamiento con el fin de describir, verificar y predecir ciertos hechos.

Esto nos lleva al lenguaje de la intelección, donde Roche hace uso de la terminología clásica para remarcar la importancia extralingüística de la terminología en la definición de conceptos a través de la combinación o factorización de características o la identificación de propiedades esenciales determinadas por la razón. Principios epistemológicos que, según Roche (Roche, 2008: 10), guían y condicionan la construcción de un sistema nocional.

La Terminología se mueve entre la lengua de uso y la lengua de la intelección, toma de la lengua de uso, los términos que observa en los textos y puede tomar algunos rasgos lingüísticos pero cuando se trata de modelar el concepto necesita más información de la encontrada en los textos (Roche, 2005: 50).

La representación de conceptos forma parte de la práctica de la Terminología y responde a dos necesidades:

- Por razones científicas, el uso de un lenguaje simbólico, claramente definido, permite disminuir los problemas de interpretación que suceden con el lenguaje natural.

- Por razones de operatividad, un sistema nocional, definido así, puede constituirse como un modelo procesable por un ordenador.

Sin embargo, Roche señala que no resulta fácil definir un formalismo ni un modelo computacional satisfactorio que pueda representar un sistema nocional (Roche, 2008: 10). Por eso, propone hacer uso de la ontología, formalización de conocimiento, para representar los sistemas nocionales e introduce el neologismo "ontoterminología" (Roche, 2008: 12). Como señalábamos anteriormente, la *ontoterminología* designa el acercamiento que sitúa la ontología en el centro de la Terminología. Un acercamiento donde la ontología juega un rol fundamental para la construcción de un sistema nocional y para la sistematización de la terminología.

Para llevar cabo el modelado de conceptos, el uso de expertos de un dominio es imprescindible. Según Roche (Roche, 2008: 12), *decir* no es lo mismo que *modelar*. El sistema lingüístico está directamente relacionado con el habla especializada pero no con un sistema conceptual, que es un tema de modelado del dominio. Cuando escribimos no estamos conceptualizando y por tanto, extraer información útil de los textos no resulta suficiente para modelar. Las estructuras léxicas (la red de términos enlazada mediante relaciones lingüísticas como la hiponimia o la sinonimia) no coincide con la estructura conceptual (entendida como una red de conceptos vinculados mediante relaciones conceptuales) y lo ilustra con el ejemplo que mostramos a continuación. En la primera figura se representa la estructura léxica de 4 términos del dominio de los relés a partir de la información obtenida en un texto, y en la segunda, la representación en forma de ontología:

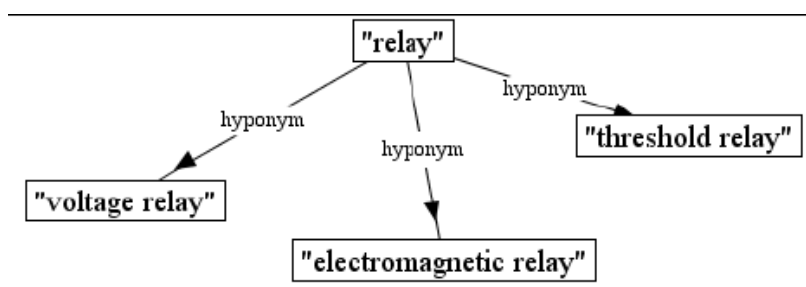


Figura 1. Ejemplo de estructura léxica de los términos derivados de *relay* (relé) (Roche 2009: 2)

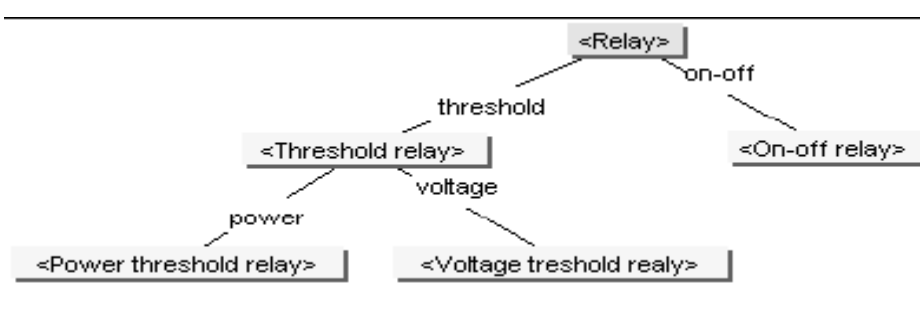


Figura 2. Ontología de *relay* (relé) (Roche, 2009: 2)

En el ejemplo, se muestra cómo en la ontología se ha representado el dominio completo de relés mientras que a partir de los textos solo se ha obtenido parte de ese dominio y su estructura no es la real. Los textos tienden a utilizar figuras retóricas como la elipsis que provocan que se anide, por ejemplo, *relé de tensión* (*voltage relay*) como un tipo de relé sin

percibir que en realidad *relé de tensión* hace referencia al concepto ‘relé de tensión umbral’ (‘voltage threshold relay’) y por tanto, es un tipo de ‘relé de umbral’ (‘threshold relay’).

En una ontología, Roche propone diferenciar entre término y concepto. Gráficamente, lo distingue introduciendo los conceptos entre signos de mayor que y menor que (<voltage threshold relay>). Además, para expresar los conceptos mediante el lenguaje propone utilizar *términos normalizados*, que son expresiones enteras del concepto que no se observan en los textos pero sirven para reflejar el concepto mediante el lenguaje. Siguiendo con el ejemplo anterior es el caso de *voltage threshold relay*, una denominación que no recibe en los textos un relé de tensión de umbral porque resulta larga pero que ayuda a entender el concepto. Estos términos normalizados se utilizan también cuando se necesita expresar un concepto que no está presente en los textos pero se considera necesario para entender el dominio que está modelando. En la Ontoterminología, los términos de uso o los normalizados están vinculados a conceptos definidos formalmente en la ontología. Esto, según Roche, facilita la gestión de las dimensiones lingüística y conceptual y proporciona dos tipos de definición: primero una formal del concepto y después una del término y uso desde un punto de vista lingüístico. Desde la ontología, los conceptos se utilizan para crear la estructura y los términos se enlazan como un tipo de información más. Eso sí, el enfoque contempla el acceso a la información tanto desde los términos como desde el concepto.

En la **estructuración conceptual**, en este enfoque al igual que en otros, se utilizan las relaciones conceptuales para elaborar las entradas y estructurarlas. Al mismo tiempo estas mismas relaciones de tipo “es-un” o “parte-de” se utilizan para construir la estructura léxica. Las palabras y las relaciones lingüísticas dejan de ser las únicas vías de acceso a la información y la información asociada a los conceptos (definiciones terminológicas, documentos, experiencias, etc.) da lugar a la clasificación del dominio experto. El resultado de una entrada se puede observar a través del siguiente ejemplo (figura 3):

The screenshot displays the OSTermin software interface. On the left, a 'Specialisation' tree lists various hydraulic components, with 'Hydraulic turbine' selected. Below this is a hierarchical diagram showing 'Hydraulic turbine' as a parent concept with sub-concepts like 'Francis hydraulic turbine', 'Pelton hydraulic turbine', and 'Turbine blade'. The main area is divided into two sections: 'Concept : <Hydraulic turbine>' and 'Terme d'usage : "turbine"'. The 'Concept' section includes a definition, composition (Hydraulic wheel, Shaft, Water box, Water inlet), and a diagram of a turbine with labels for Generator, Turbine, Turbine shaft, Water inlet, and Turbine blades. The 'Terme d'usage' section provides a linguistic definition and a list of synonyms: hydraulic wheel, shaft, water inlet, water box, and wheel.

Figura 3. Extracto de una ontoterminología de turbinas hidráulicas y su término en francés (Roche, 2009: 4).

En el ejemplo, vemos como la estructura se basa en los conceptos que a su vez están vinculados a información lingüística (término, definición, etc.). Además, en la ventana derecha superior, podemos ver la distinción entre término de uso y término normalizado y como ambos conectan con el mismo concepto. Esto mismo se puede realizar en más de una lengua, en ese caso, el término normalizado seguirá siendo solo uno y los términos de uso serán tantos como lenguas o sinónimos en una misma lengua queramos incluir.

La diferenciación y estructuración de los conceptos del dominio se basa en el principio de **diferencia específica** y consiste en identificar las características eidéticas de éstos dentro del sistema con el fin de distinguirlos según las que no comparten. Una característica eidética es aquella que en su ausencia haría que el objeto ya no fuera tal objeto. La definición de los conceptos en esta teoría se basará también en eso, en lo que lo hace diferente con respecto al resto en el sistema. El concepto se observa en este caso como un conjunto de atributos y mediante estos se define y diferencia a un concepto de otro (Roche, 2001). Mediante este principio, el autor defiende que las definiciones serán consistentes desde un punto de vista lógico y permiten alcanzar un consenso entre distintos expertos y/o terminólogos ya que el debate no se centrará en la definición de un concepto concreto sino en la identificación de lo que lo diferencia de otros, algo más específico y por tanto, más sencillo de consensuar.

Para organizar todos estos conceptos y sus descripciones, la Ontoterminología hace uso de **esquemas**, imitando los que se utilizan en ingeniería o en la industria para describir los objetos del dominio. En su enfoque, un esquema describe una entidad física y las partes que la conforman. Cada una de estas partes también se describe mediante un esquema de forma que las entidades y sus componentes están enlazadas mediante relaciones de tipo “parte-de”. Los conceptos crean una red enlazada mediante relaciones “parte-de”, que permiten al usuario acceder de un esquema a otro más detallado o más global. Esto se realiza mediante un método de navegación que está vinculado a la ontología.

En conclusión, la **ontoterminología** insiste en la importancia de los principios epistemológicos que preceden a la conceptualización de un dominio y en la necesidad de un acercamiento científico donde la ingeniería juega un papel fundamental.

Esta propuesta se justifica en que hay dominios científicos y técnicas que requieren la conceptualización del mundo y la creación de denominaciones unívocas para sus constituyentes, por ejemplo, en ingeniería colaborativa, donde distintas comunidades de especialistas forman parte de la concepción o fabricación de un producto. Por esto, Roche sugiere que siendo la terminología una disciplina autónoma que toma aspectos de la epistemología, la lógica y la lingüística, puede suceder que la terminología se centre en el concepto y no tanto en aspectos lingüísticos.

1.6 Recapitulación y observaciones

La terminología es una disciplina que estudia y recopila los términos de un ámbito de especialidad para facilitar la comunicación. Como los términos se pueden encontrar en los textos especializados, se hace uso de la lingüística para extraerlos y estudiarlos, por ejemplo, mediante el análisis textual. Esta tarea no exime al terminólogo de hacer uso de otro tipo de razonamientos, de carácter cognitivo o comunicativo, cuando trata de estructurar, describir y

representar la terminología de un ámbito. Este razonamiento se podría amparar en la teoría de Cabré, que remarca que se puede acceder a los términos desde diferentes planos: el lingüístico, el cognitivo y el comunicativo.

A continuación, analizamos las aportaciones de Meyer, Temmerman, Faber y Roche al propósito de este trabajo.

Las **unidades de estudio** de estas cuatro teorías se centran en los términos y en el caso de Temmerman, en las *unidades de comprensión* por diferenciación de los *conceptos*. Faber también analiza otro tipo de unidad además de conceptos y términos, que tiene lugar en la representación de conceptos y denomina *categorías*. Roche añade al estudio de conceptos y términos, el estudio de la *ontoterminología* (como representación formal de una conceptualización) y lo sitúa como el centro de interés de su teoría.

En cuanto al **objetivo** de estas cuatro teorías, la aportación de Meyer se centra en desarrollar herramientas que ayuden a facilitar el trabajo de representación de conceptos y términos. En el caso de Temmerman, su propuesta va dirigida a la creación de ontologías teniendo en cuenta la información terminológica de un dominio y sus particularidades socio-culturales. Faber se centra en la creación de sistemas conceptuales que permitan representar su naturaleza multidimensional y su uso en el lenguaje a través de términos. Roche parece coincidir con Faber en este aspecto y propone el uso de ontoterminologías, donde los conceptos y los términos se definen de manera independiente. Además, Roche propone distinguir entre la representación del conocimiento a través de una ontología y la representación de información basada en el análisis lingüístico de los textos. Forma parte de este trabajo que presentamos profundizar en este aspecto y en las estrategias de formalización de estos dos enfoques.

Un factor común en estas teorías es la reivindicación de la **dimensión conceptual** con respecto a la dimensión lingüística en el estudio y representación de los términos. En la Ontoterminología de Roche se enfatiza un enfoque que parte de los conceptos del dominio para llegar a la terminología de uso. Un aspecto interesante de este enfoque es el uso de *términos normalizados* para hacer referencia a conceptos que no tienen un término de uso en el lenguaje o cuya denominación resulta poco transparente. Esto evidencia la posición prioritaria del concepto con respecto al término en su planteamiento. Temmerman también tiene su aportación en este aspecto cuando señala que los conceptos no solo lo son por existir en la realidad sino por poder ser concebidos en la mente.

Además, todas proponen el uso de un nuevo recurso terminológico distinto al diccionario, que varía según la teoría utilizada (base de datos de conocimiento terminológico a ontología u ontoterminología) pero que comparte una base común y es la representación formal de conceptos y sus relaciones.

Para iniciar la **estructuración de conceptos** y las relaciones entre conceptos, se propone el uso de redes conceptuales en el caso de Meyer o de marcos semánticos basados en los eventos del dominio en el caso de Faber. Temmerman defiende el desarrollo de modelos cognitivos y en cambio Roche se acerca más a la metodología de los ingenieros informáticos y propone la elaboración de esquemas conceptuales de las entidades que forman parte del dominio a describir. Aun así, notamos diferencias en los criterios para estructurar los conceptos. Por ejemplo, Temmerman defiende que la estructuración de conceptos no se puede basar únicamente en un análisis de sus características esenciales y en cambio Roche propone el principio de diferencia específica como único criterio para diferenciar los conceptos. Es este

aspecto, consideramos muy interesante investigar más en detalle esta propuesta de Roche y su compatibilidad con la representación de la multidimensionalidad de conceptos.

2 Las características

2.1 El término

Felber y Picht en su libro sobre métodos de terminografía y principios de investigación terminológica (Felber y Picht, 1984: 211) se acogen a la definición de terminología que proponía la norma ISO 1087 de 1969 (ISO, 1969), con la que coincide la norma UNE-EN 1070 de 1979 (AENOR, 1979), y consideran ‘término’ a la representación lingüística de un concepto. Sin embargo, más tarde Picht en su trabajo con Arntz de 1995 (Arntz y Picht, 1995: 58) reconsidera el papel de ‘término’ al observarlo como una unidad constituida por un concepto y una denominación. Este nuevo acercamiento al término será el que prevalezca en la mayoría de trabajos sobre terminología aunque las normas internacionales de estandarización continúen reflejando el carácter de término desde una perspectiva lingüística.

L’Homme et al. (L’Homme et al., 2003: 154-157) revisan la definición de término y las distintas aportaciones que han intentado establecer o identificar las posibles diferencias o parecidos entre un término y una palabra. Desde un punto de vista lingüístico, no se detectan diferencias notables como demuestran distintos estudios destinados y se corrobora que existe una línea muy difusa entre término y palabra desde un punto de vista lingüístico (Kageura, 1997, 2002; Meyer y Macintosh, 2000).

Para distinguir entre término y palabra se recurre entonces al dominio o al entorno extralingüístico en el que tiene lugar el uso de un término. En el primer caso se suelen asociar a su uso en textos que se compilan en forma de corpus con fines específicos.

Respecto a lo que representa un término, en la TGT se entiende que el término representa a un concepto, mientras que las palabras que hacen referencia a nociones. Esta idea parece reflejarse en la definición de término de la norma UNE-EN 1066 (AENOR, 1991: 9): «una palabra o grupo de palabras utilizado para designar un concepto».

En la norma UNE-EN 1070 (AENOR, 1979: 7), reemplazada por la UNE-EN ISO 1087 (AENOR, 2009), se definía término de una noción (concepto) como:

«El símbolo convencional de una noción que consiste en sonidos articulados o en su representación gráfica. Un término es una palabra o un grupo de palabras»

En esta definición no se aludía a la relación entre los términos y los ámbitos de especialidad pero se complementaba con una referencia a lo que denominan “término técnico”: «el término cuyo empleo o el sentido está restringido a los especialistas de un campo particular del saber».

La norma UNE-EN ISO 1087 (AENOR, 2009: 9), que define ‘término’ como «Designación verbal de un concepto general en un campo temático», se acerca al término desde enfoques más modernos y con un límite más difuso entre término y palabra. En el análisis de L’Homme et al. (ibíd.), se presentan otras aportaciones que consideran los términos como meras unidades léxicas portadoras de significado especializado. Es el caso de Kocourek, (Kocourek,

1991) que entiende que el significado de un término es como el significado de una palabra, solo que en el caso de los términos algunos significados solo se dan en un dominio especializado. Otro punto de vista presentado es el de Cabré (Cabré, 1992) que entiende que el significado especializado se activa cuando tienen lugar unas circunstancias comunicativas concretas. Temmerman (Temmerman, 2000: 223) entiende los términos como una unidad de significado. En este aspecto, L'Homme y otros coinciden con Cabré (ibíd.) en afirmar que los términos se pueden observar como poliedros (unidades multidimensionales) que pueden observarse desde distintos ángulos según las necesidades y el objetivo de la tarea. Visto así, se entiende que convivan numerosos estudios sobre el término, y que éstos se centren en distintas dimensiones del término (Sager, 1993): la vertiente lingüística del término (su formación o su identificación en los textos), la vertiente cognitiva (su definición y representación) o la vertiente comunicativa (usuarios y dimensión social).

2.1.1 El concepto

Un concepto se define según la norma UNE-EN ISO 1087 (AENOR, 2009: 6) como: «Unidad de pensamiento constituida por aquellas características que se atribuyen a un objeto o una clase de objetos».

Esta norma distingue entre conceptos únicos y conceptos generales. Un 'concepto general' se define como: «un concepto que corresponde a dos o más objetos que forman un grupo debido a las propiedades que tienen en común» (ibídem: 7).

Un concepto único es aquel que corresponde a un único objeto, por ejemplo, 'Saturno' o 'la torre Eiffel' (ibídem: 6).

Un objeto es «todo lo que puede ser percibido y conocido» (AENOR, 2009: 6), y puede ser material, inmaterial o imaginario.

En la bibliografía sobre Terminología, existen muchas otras definiciones de concepto. Pitch y Draskau (Picht y Draskau, 1985) estudian dichas definiciones y explican que si bien la definición de concepto puede variar según la escuela y el área que la formula, las aportaciones de cada escuela determinadas por el área que estudian se han demostrado aplicables a otras áreas en muchos casos.

La definición de concepto de Wüster se acerca mucho a la ofrecida en la UNE-ISO 1087 (AENOR, 2009) y habla de que los conceptos, con excepción de los conceptos individuales, corresponden a elementos comunes que los seres humanos perciben en un gran número de objetos y que utilizan como medio de clasificación mental y para comunicarse. Con lo que argumenta que un concepto es un elemento del pensamiento. (Wüster 1998: 39)

La norma DIN 2330 (DIN, 1979) coincide con la reflexión de Wüster aunque se toma como punto de partida el cálculo de predicados de la lógica formal, contrario a la propuesta de Wüster que parte del análisis del concepto. En realidad, como demuestran Picht y Draskau (ibídem) las dos definiciones no muestran conflicto sino que funcionan en paralelo hacia un mismo objetivo. En el caso de Wüster, cabe decir que las características se extraen a partir de un análisis y en el caso del trabajo de Dahlberg (Dahlberg, 1976) los predicados del concepto llegan a establecer las características.

En Cabré (Cabré, 1992: 181), el concepto se presenta como una unidad del plano del contenido integrada por un conjunto sistemático de características de distintos tipos, lo que

añade una dimensión más detallada sin dejar de coincidir con las propuestas de Wüster o Dahlberg. Coincidiendo con lo establecido en la norma UNE-ISO 1087 (AENOR, 2009), Cabré señala que cada una de las propiedades que lo describen y la expresión del conjunto de las características de un concepto es su paráfrasis o definición.

Kageura (Kageura, 2002: 22-24) acerca el concepto al significado para poder explicarlo y señala que la definición de concepto como unidad de pensamiento, que consiste en un conjunto de características que a su vez son conceptos, es en esencia la misma caracterización de *aquello* que se entiende desde un punto de vista descriptivo como *significado*. Considerando que, en lexicología, el significado es un conjunto de características semánticas (Lyons, 1977), Kageura apunta hacia el paralelismo entre esta disciplina y las definiciones aportadas por la TGT para el concepto y plantea una alternativa que se aproxima a la categorización de términos. Esta aproximación tiene como base que la terminología de un dominio, vista independientemente, puede observarse como poseedora de una estructura propia y, por tanto, como representante de los conceptos de un dominio. El concepto se observa desde la lingüística, como un significado restringido y claro sobre un dominio. Este enfoque, orientado a la terminología aplicada y tomado de Sager (Sager, 1993), plantea la no necesidad de definir concepto ya que este es una primitiva axiomática más como lo son “frase” o “palabra” con las que trabaja el terminólogo.

Desde un enfoque cognitivo, la teoría de conceptos para la Terminología solo tiene que proporcionar una explicación adecuada de las motivaciones cognitivas en la formación de términos y proporcionar una base para estructurar vocabularios de forma efectiva (ibídem: 15). Siendo los términos la materia prima de la terminología, que se conciben como los símbolos que representan conceptos.

2.1.2 La intensión y extensión de un concepto

Intensión y extensión se consideran conceptos centrales en la teoría de la terminología y son de especial interés para la formación de conceptos y para la teoría de definiciones. Según la definición de la norma UNE-ISO 1087 (AENOR, 2009), la *intensión* es el conjunto de todas las características que constituyen un concepto y sirve a la mutua delimitación de conceptos. Un ejemplo simplificado de la intensión de lobo sería “mamífero, carnívoro, fisípedo y cánido” (Cabré, 1992: 186).

En cuanto a la definición de *extensión*, la norma ISO 1087 de 1969 (ISO, 1969) contemplaba dos significados. Por un lado, desde el punto de vista de similitud, como el conjunto de todas las especies imaginables de un concepto considerado por separado, ocupando todas las especies el mismo nivel de abstracción respecto al objeto; por ejemplo, la extensión del concepto elefante incluiría elefante indio y elefante africano. Por otro, desde el punto de vista de su composición, como el conjunto de partes de un todo, consideradas por separado. Esta última acepción fue cuestionada por distintos autores, ya que no se consideraba estrictamente la extensión de un concepto sino la parte funcional o la suma de los componentes que lo forman (Picht y Draskau, 1985).

Si consideramos el papel que juegan las características del concepto en este entorno, notaremos que su presencia está relacionada con el grado de extensión e intensión que tenga un concepto. Es decir, a mayor extensión menos características distintivas poseerá un concepto, ya que el número de objetos incluidos será mayor y por tanto su especificidad menor. En cambio, cuanto mayor sea el grado de especialización de conceptos, es decir, cuantas más características lo distinguen del resto de conceptos que pertenecen a su misma clase en un sistema de conceptos,

la extensión y el número de objetos serán menores. Este razonamiento se expresa en la norma UNE 1066 (AENOR, 1991) como:

«A más extensa sea la comprensión de un concepto, es decir, cuanto mayor sea el número de características que incluya, más restringida será su extensión y menor el número de objetos que poseen dichas características. Por el contrario, al restringir la comprensión aumenta la extensión.» (AENOR, 1991:3).

Esto mismo vienen a decir Meyer et al. (Meyer et al., 1997: 101): un conjunto de características representa el conocimiento sobre un concepto y a más características, más conocimiento se representa sobre un concepto. La suma de todas las características se denomina *intensión*, y la colección de objetos a los cuales hace referencia un concepto se denomina *extensión*. Un concepto que tiene muchas características se dice que tendrá una *intensión amplia* y su *extensión* será corta.

2.1.3 Tipos de concepto

Aunque no se ha llegado a un acuerdo para explicar la formación de conceptos, esto no ha impedido su análisis y por tanto estudiar las características ligadas a cada concepto, que resultarán de gran interés en Terminología.

Cuando un objeto nuevo se crea o se descubre, aparece un nuevo concepto, cuyo tipo y configuración en cuanto a sus características requiere un análisis con la ayuda de predicados. La suma de predicados (que establecen las características) es igual a la suma de nuestro conocimiento sobre un concepto (Picht y Draskau, 1985).

Los dos tipos de conceptos más admitidos por la mayoría de autores son: conceptos individuales y conceptos generales. La presencia del primer tipo de elementos, los conceptos individuales, siempre produce un concepto individual pero que el segundo tipo estará presente en ambos grupos.

Los conceptos individuales representan objetos individuales y los conceptos generales, objetos indefinidos. En cuanto a la facilidad para distinguir un objeto individual de uno general, basta con intentar situar el objeto en un espacio y un tiempo determinados, si esto es posible se tratará de un objeto individual. No obstante, el fenómeno de la individualización dificulta su reconocimiento en cierto modo, esto ocurre cuando a un objeto general se le añade un componente espacio-temporal (por ejemplo, mi baldosa [la que tengo en mis manos en este momento]).

Además de hacer referencia a objetos individuales y objetos generales, necesitamos poder estructurarlos para poder representarlos y trabajar con ellos. La clasificación de conceptos servirá para entender mejor un dominio específico y representar todo el conocimiento de forma estructurada y coherente. Un modelo de agrupación desde un punto de vista cognitivo es la que propone Sager (Sager, 1993: 25) que distingue 4 métodos que devuelven 4 grandes grupos de conceptos:

- Los conceptos que se pueden agrupar en categorías establecidas por las características distintas de cada clase y por grupos de clases.
- Los conceptos que se pueden atribuir a una clase, de modo que se identifican tipos de concepto.

- Los conceptos y grupos de conceptos que se pueden distinguir mediante un proceso de discriminación entre categorías y mediante relaciones.
- La interacción de categorías de conceptos que se puede reflejar a nivel funcional, para determinar estados y cambios de estado.

Estos puntos sirven para obtener conceptos de clases, propiedades, relaciones y funciones que variarán dependiendo de la situación y del ámbito. Ejemplos de tipos de concepto serían contenedor (clase), albino (propiedad), paralelo (relación) y operación (función). Dependiendo del nivel de detalle en el que se incluyan los conceptos se habla de niveles superiores (los más generalizables) o niveles inferiores (los más detallados). Esta clasificación de clases, propiedades, relaciones y funciones coincide con la mayoría de las propuestas para representación del lenguaje mediante sistemas computacionales (Pugh, 1984; Carlson y Nirenburg, 1990; Fellbaum, 1998).

Otro tipo de clasificación viene determinado por la asociación lingüística a los conceptos, de donde se produce una diferenciación entre conceptos-adjetivo, conceptos-verbo y conceptos-nombre. No obstante, normalmente se han venido utilizando solamente los nombres para representar los conceptos.

2.1.4 **Definición de característica**

Por definición, la norma UNE 1066 define las características desde su vertiente funcional:

«Las características sirven de base para la clasificación de conceptos. Son necesarias para diferenciar un concepto de otro dentro de un área específica así como también para otras funciones» (AENOR, 1991: 3).

Esta necesidad para diferenciar conceptos también se ve reflejada en la norma DIN 2330 (DIN, 1979: 6) donde se explica que las características de los conceptos tienen una importancia especial tanto para la definición del concepto como para la determinación de las relaciones entre conceptos. En esta norma se explica que las características se basan en las propiedades de los objetos determinadas, por ejemplo, mediante la observación, resultados de medición, aseveraciones generalmente aceptadas sobre objetos o estipulaciones normalizadas. Además, añade que las propiedades que deban considerarse como características han de seleccionarse con sumo cuidado para que queden claramente establecidas y se puedan reconocer con facilidad. Arntz y Picht (Arntz y Picht, 1995: 79) puntualizan que esta selección se realiza con un fin pero que las características están presentes o ausentes en un concepto.

Al igual que en el trabajo de Arntz y Picht (Arntz y Picht, 1995: 173), Kageura cuando trata las características remarca que por una parte, una característica es, para un sistema de conceptos, una propiedad de un concepto, la faceta más importante para la descripción de conceptos; pero por otra, como elementos que constituyen parte de un sistema de conceptos, las características también son conceptos en sí mismos y por tanto, se explica el hecho de que una característica lo sea en un caso y que, en cambio, en otro, sea observada meramente como un concepto (Kageura, 2002: 67). Más adelante, veremos cómo se refleja esto su clasificación de tipos de características (§2.2.4).

Un aspecto que no podemos dejar de mencionar es el que refleja la importancia de las características para el trabajo terminográfico. Como veremos en los cuatro criterios presentados

a continuación, establecidos por Arntz y Picht (Arntz y Picht, 1995: 79), las características jugarán un papel determinante en la terminología aplicada:

- *Determinación de la comprensión de un concepto.* Ya que la totalidad de características determinadas en un momento dado constituye la suma de conocimiento acerca del mismo, la modificación de una característica daría lugar a un concepto diferente.
- *Fundamento de la formación de términos.* Cuando se crean términos nuevos o se analizan o evalúan términos existentes, las características que han de quedar expresadas lingüísticamente tienen una especial importancia. No obstante, la subsistencia de un término no es un indicador seguro de que el concepto no se haya modificado, como es el caso del término *átomo*.
- *Estructuración de sistemas de conceptos.* Los tipos de características tienen una influencia decisiva en la estructura de un sistema de conceptos, en las relaciones sistemáticas entre conceptos. Como criterio de clasificación, las características determinan qué conceptos deberán aparecer directamente unos al lado de otros o en un orden jerárquico.
- *Determinación de equivalencias.* Si los conceptos representados por unas denominaciones coinciden en sus características (identidad conceptual), las denominaciones son sinónimas (en una misma lengua) o equivalentes (en lenguas distintas).

A partir de esto, se confirma la opinión generalizada de que los conceptos se distinguen porque poseen características diferentes y por tanto, una descripción conceptual adecuada deberá recoger la contraposición de rasgos con capacidad distintiva que diferencia los conceptos entre sí, como señala Cabré (Cabré, 1992: 182).

«La presència (o absència) d'una característica en lloc d'una altra en un conjunt de característiques interrelacionades dóna raó de la diferència entre els conceptes d'un mateix sistema conceptual» (Cabré 1992: 183).

Por tanto, hay características comunes que hacen que un grupo de conceptos se entienda como perteneciente a un mismo campo y características en relación de oposición que los hacen diferentes entre ellos. Al mismo tiempo, Cabré señala que una característica puede ser descrita y formar parte de un diccionario.

En el trabajo de Madsen et al. (Madsen, 1998; Madsen et al., 2005) las características juegan un papel muy importante ya que son la base del criterio de subdivisión de conceptos y además las responsables de representar la multidimensionalidad de un sistema conceptual. *Característica* se define:

«an element of a concept denoting a property of an entity or a class of entities» (Madsen 1998: 342)

y se adhieren, además, a la definición proporcionada por Nuopponen, donde una característica se define como:

«a concept, the referent of which is a property of an entity» (Nuopponen, 1994: 61).

Estas dos definiciones, según Madsen, no son contradictorias sino que funcionan en paralelo y representan, por un lado, el carácter de elemento que constituye un concepto y por otro lado, el de concepto en sí mismo (Madsen, 1998: 342). Además de esto, Madsen y otros señalan que las relaciones conceptuales son características vistas desde un enfoque de representación de un sistema conceptual y en su trabajo las representan como tales (Madsen et al., 2005: 162). Según su enfoque, no existe distinción entre describir un concepto mediante características, por ejemplo, adscribir a ‘book’ (‘libro’) la característica ‘has a spine’ (‘tiene lomo’), o hacerlo mediante relaciones con otros conceptos, con el ejemplo de una relación parte-todo entre ‘book’ (‘libro’) y ‘spine’ (‘lomo’), ya que consideran que son dos formas de expresar lo mismo.

Desde el punto de vista de la clasificación de conceptos, las características juegan un papel fundamental, ya que una clase de conceptos es un grupo de conceptos que tiene al menos una característica en común (Bowker, 1997: 133). Se dice que cuando se puede utilizar más de una característica para distinguir conceptos, permitiendo distintas clasificaciones, hay distintas dimensiones. Cada dimensión representará una forma de clasificar un grupo de conceptos, una clase con más de una dimensión se denomina *multidimensional*. Por ejemplo, los tipos de ‘vehículo’ se pueden clasificar según tres características:

- el medio de transporte (vehículo acuático, vehículo terrestre y vehículo aéreo)
- el método de propulsión (vehículo motorizado, vehículo no motorizado)
- el tipo de carga que transportan (vehículo de transporte de pasajeros, vehículo de transporte de pasajeros y materiales, vehículo de transporte de materiales)

Según este ejemplo de Bowker (ibídem: 135), un vehículo tiene tres dimensiones, es decir, se puede agrupar atendiendo a tres características. Así, el concepto ‘canoa’ pertenecerá al medio de transporte acuático pero también se puede agrupar como vehículo no motorizado o vehículo capaz de transportar a personas y materiales.

Este fenómeno, conocido como la **multidimensionalidad del concepto**, se ilustra claramente cuando el concepto se observa como un conjunto de características, es decir desde un enfoque intensional. La organización estática del sistema de conceptos se puede observar como la totalidad de todos los tipos de características que se utilizan para constituir todos los conceptos en el sistema. Entonces, este sistema, estático, porque se ve como una representación de conceptos establecidos por términos y sus inter-relaciones, se muestra como un espacio multidimensional donde los ejes de intersección representan cada faceta o tipo de característica según la cual se especifica cada característica de un concepto. Este enfoque ha sido desarrollado en profundidad por Kageura (Kageura, 1997: 120) y resulta muy similar al propuesto por Sager (Sager, 1993) dentro de su modelo de conocimiento. En la siguiente figura, encontramos la representación gráfica de este planteamiento del espacio multidimensional del concepto representado en el trabajo de Kageura:

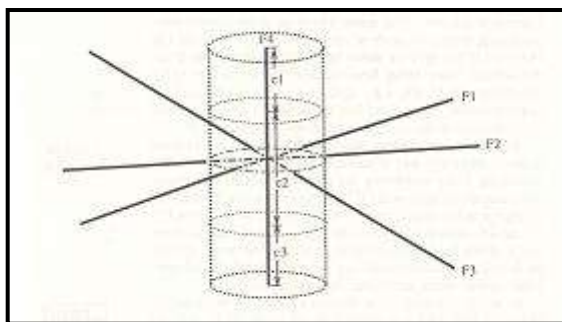


Figura 4. Espacio multidimensional en un sistema de conceptos (Kageura, 1997: 121)

En esta figura, C1, C2 y C3 situados a lo largo de F4 representan los rangos del tipo de característica F4, es decir los posibles valores para la característica. La intersección de los rangos F1-F3 situada en el centro del dibujo y delimitada por unos puntos suspensivos determina el sub-espacio que abarcaría un concepto hipotético. Por tanto, en esta figura, se representa al concepto constituido por las características F1, F2, F3 y F4, y como ejemplo se han dibujado los posibles rangos de F4, uno de los cuales caracteriza al concepto, aunque debemos imaginar, que ocurre lo mismo con las otras tres características, cuyos rangos no se han dibujado para representar este espacio multidimensional. Esta representación intenta simular un espacio en tres dimensiones donde las distintas coordenadas sirven para delimitar un concepto dentro de su sistema, visto desde un enfoque intensional, es decir, observando las características que lo constituyen.

2.1.5 Recapitulación y observaciones

En este capítulo pretendíamos establecer las nociones sobre el significado de ‘característica’ y reflejar, a través de las distintas definiciones, la importancia de las características para la Terminología. En este trabajo tratamos de describir las características de un concepto de un ámbito de especialidad. Los conceptos son un elemento abstracto y coincidimos con la mayoría de teorías que abogan por abarcarlos desde sus características, que sirven para determinar su comprensión, para estructurarlos y en definitiva, para describirlos.

En cuanto a la clasificación de tipos de concepto, consideramos que la clasificación de Sager, en clases, propiedades, relaciones y funciones distingue de una manera clara entre propiedades y relaciones.

En cuanto a su papel en la descripción de conceptos de las características, entendemos que en las propuestas de Madsen o Nuoponnen la definición de características está enfocada hacia la descripción de entidades, de ahí que se obvie aludir al papel de las características como descriptoras del resto de conceptos.

Coincidimos con la aproximación a las relaciones de Madsen et al., cuando mantienen que no existe distinción entre describir un concepto mediante características o hacerlo mediante relaciones con otros conceptos. De hecho, creemos que esta distinción en el plano de la representación de conceptos dependerá de lo que queramos reflejar en la representación. Por ejemplo, si lo que interesa para una ontología es reflejar un rasgo de un concepto, conviene representar ese rasgo en forma de característica, en cambio si lo que queremos es hacer hincapié en la relación que se mantiene entre dos conceptos, convendrá resaltar que existe tal relación.

Para identificar las características de un concepto buscamos en los textos especializados pistas a través de los términos y otras construcciones presentes. Esto se fundamenta en tres ideas básicas:

- 1) El término es la unidad de estudio de la Terminología, éste se compone de denominación y concepto.
- 2) Los términos se pueden encontrar en los textos representados por sus denominaciones.
- 3) Los conceptos se pueden clasificar de forma distinta según las necesidades de la clasificación.

En lo que se refiere a la multidimensionalidad del concepto, hemos incluido este apartado aquí con el fin de hacer ver dos cosas: por una parte, la importancia de las características en la descripción de las distintas dimensiones del concepto, y por otra, que esta multidimensionalidad, las distintas formas de clasificar, aparece cuando describimos una característica al ser vista como un concepto.

2.2 Tipos y partes de características

Cuando analizamos la bibliografía sobre características encontramos clasificaciones que están presentes en casi todas las obras:

- Características esenciales y no esenciales (Felber y Picht, 1984; Sager, 1993; Cabré, 1992; Arntz y Picht, 1995)
- Características intrínsecas y extrínsecas (Felber y Picht, 1984; Sager, 1993; Cabré, 1992; Arntz y Picht, 1995).
- Características atributo y valor (Meyer et al, 1997; Kageura, 2002; Madsen, 1998 y Thomsen, 1998).

A éstas clasificaciones se suman otras más particulares que cada autor ha utilizado para organizar las características de acuerdo con sus propósitos de investigación y que agrupan las características a partir de otras más generales (Madsen, 1998; Kageura, 2002; Vaupel, 2005). Además, encontramos otras clasificaciones que si bien no aparecen formalmente en las tipologías sobre características se utilizan en la terminología aplicada y pueden servir para clasificar características. En este apartado, analizamos algunos de estos estudios con el fin de identificar qué aspectos influyen en la descripción de una característica.

2.2.1 Características esenciales, distintivas y no esenciales

Felber y Picht señalan que:

«Las características que se emplean en un proceso específico del análisis conceptual, para la definición o determinación de las relaciones entre conceptos, o sea, para ordenar conceptos, se llaman *características de orden*. Las características restantes se consideran como no fundamentales» (Felber y Picht, 1984: 176)

Estos autores además puntualizan que al cambiar el aspecto o la orientación es posible que las características no fundamentales se conviertan en características de orden y viceversa (ibíd.).

Actualmente, en trabajos más recientes sobre terminología como el de Cabré (Cabré, 1992: 184) o Sager (Sager, 1993: 48) y en la norma UNE 1066 (AENOR, 1991) encontramos una clasificación similar bajo la denominación de características esenciales y no esenciales. Según Cabré:

«las características esenciales de un concepto son las que describen su esencia y, por tanto, son ineludibles para definirlo». (Cabré, 1992: 183)

En palabras de Sager (Sager, 1993: 48), las características esenciales son aquellas que resultan adecuadas y necesarias para identificar un concepto. Siguiendo el ejemplo de Sager (Sager, 1993: 48) las características esenciales de ‘mesa’ serían:

- horizontal
- de superficie lisa
- dentro de unos límites dimensionales y de proporción de las dimensiones
- de una alzada sobre el suelo no superior a un metro
- utilizada para colocar objetos o trabajar sobre ella

Las características esenciales se oponen a las características no esenciales o complementarias. Las características no esenciales son las que añaden elementos no relevantes para la descripción de un concepto (Cabré, 1992: 184). En Cabré (ibíd.), a modo de ejemplo, se indica que las características esenciales de ‘poliamida’ son: conjunto de compuestos poliméricos obtenidos por reacción entre dianhídridos y diaminas.

En palabras de Sager, las características no esenciales son aquellas que no resultan suficientes para distinguir un concepto. Siguiendo el ejemplo de Sager (Sager, 1993: 48) las características no esenciales de ‘mesa’ serían: utilizada para colocar objetos o trabajar sobre ella, color, tipo de material.

Tanto Sager (Sager, 1993: 48) como Cabré (Cabré, 1992: 184) coinciden con Felber y Picht en afirmar que las características no esenciales para un concepto pueden convertirse en esenciales para la creación de otros conceptos en otros planos de especialidad y asimismo, las no esenciales pueden resultar esenciales dependiendo del plano de especialidad. En su ejemplo, Sager (ibídem: 49), explica que al fijar la atención en distinguir entre ‘mesilla de noche’, ‘mesilla de café’ y ‘mesa de comedor’, las características esenciales referirán también a la altura y el tamaño, pero serán más específicas y por tanto, alzada sobre el suelo no superior a un metro dejará de ser relevante.

Además, entre las características esenciales, la norma ISO 704 distingue un subgrupo de características denominadas distintivas. La norma señala que cuando estudiamos un concepto, cada una de sus características esenciales debe estudiarse en relación a otros conceptos vinculados al sistema de conceptos. Dentro de este análisis, «*un caractère distinctif est un caractère essentiel qui permet de délimiter un concept par rapport à un autre*» (ISO, 2000: 5). Las características distintivas se entienden como un tipo de característica esencial y podrán ser distintivas para un concepto y no para otros y viceversa.

Madsen et al. (Madsen, 1998; Thomsen, 1998; Madsen et al., 2005) se cuestionan la necesidad de diferenciar entre características esenciales, distintivas y no esenciales porque

consideran que las características que no resultan esenciales para describir un concepto desde un determinado ámbito no deben aparecer en la descripción del concepto. En su lugar, proponen distinguir entre características delimitadoras y no delimitadoras. Este binomio considera que las características se deben estudiar siempre en su marco conceptual, es decir, observando las propiedades que tiene un concepto como miembro de un grupo de conceptos.

Las *características delimitadoras* serán aquellas que distinguen un concepto del resto de conceptos situados en un mismo nivel de abstracción y clasificados bajo un mismo criterio. El resto de características que se consideran necesarias pero no suficientes para definir el concepto en ese nivel de abstracción, se denominan *características no delimitadoras*. Esta propuesta no está enfocada a la formalización de todas las características de un concepto, sino que sirve únicamente a propósitos de distinción entre conceptos en un determinado sistema (Madsen et al., 2005: 162, véase también §3.1.4.2).

2.2.2 Nombres de característica y valores de características

Otra forma de diferenciar tipos de característica se basa en el papel que juegan en la representación de conceptos. Meyer explica que los conceptos pueden describirse utilizando lo que en terminología se denomina *características* y en ingeniería de conocimiento *atributos* y *propiedades* (Meyer, 1997: 100). Este enfoque cuestiona que las características sean una entidad única y explica que de cara a su codificación, puede resultar más interesante considerarlas una entidad de dos componentes: el nombre de la característica y el valor de la característica. Además, Meyer señala que tanto los nombres de una característica como los valores de una característica pueden ser conceptos y pueden denotar conceptos especializados.

Madsen y Thomsen (Madsen, 1998; Thomsen, 1998) desarrollan este aspecto con el fin de formalizar las características de los dispositivos médicos y señalan que las características consisten, en realidad, en una relación entre el concepto a definir y otro concepto asociado que sirve para definirlo (Thomsen, 1998: 356). Esta relación se representa mediante un atributo y el concepto asociado será el valor del atributo, por ejemplo el par “*coverage_period: lifetime*” representa la característica que se compone del atributo “*coverage_period*” y el valor “*lifetime*”.

Esta distinción entre dos tipos de característica: nombres y valores aparece también en el trabajo de Kageura (Kageura, 2002). En su trabajo, Kageura (Kageura, 2002: 77) señala que las cualidades operan en dos niveles de abstracción al mismo tiempo. En un primer nivel de abstracción se identifican los valores de las cualidades, que son concretos; y en un segundo, se encuentran las etiquetas o rangos de dichos valores y grupos de valores, que corresponden a etiquetas abstractas de los valores. Según este criterio, 'color' será una etiqueta o nombre de una característica y 'rojo' un valor. Este binomio recibe distintas denominaciones, por ejemplo, encontramos *nombre: valor* en trabajos como el de Meyer (Meyer, 1997: 101) o *atributo: valor* en el de Madsen et al. (Madsen et al., 2005: 162) o *etiqueta: valor* en el de Kageura (Kageura, 1997: 77).

2.2.3 Características intrínsecas y extrínsecas

Distintos autores hacen referencia a la distinción entre características *intrínsecas* y *extrínsecas* (Wüster, 1979; Picht y Draskau, 1985; Sager, 1993; Cabré, 1992).

Wüster distingue entre características intrínsecas o inherentes y características de extrínsecas o relacionales. Según Wüster, (Wüster, 1998: 46-47):

- Una característica intrínseca (o inherente) es una característica que se puede observar mediante un simple examen del objeto y que no requiere más conocimientos sobre el uso o el origen de dicho objeto.
- Una característica extrínseca (o relacional) es una característica que describe la relación del objeto con otros objetos.

Dentro de las características extrínsecas, Wüster distingue dos grupos:

- Características de uso
- Características de origen

En esta aproximación, Wüster entiende que la distinción entre características extrínsecas o intrínsecas se aplica a objetos o productos materiales. Sin embargo, desde nuestro punto de vista, esta clasificación puede aplicarse también a las características que describen objetos abstractos.

Cabré (Cabré, 1992: 185) define las características intrínsecas y extrínsecas de la siguiente forma:

- Las características intrínsecas son aquellas que están ligadas a su descripción como representante de una clase; por ejemplo, forma, color o medida en el caso de los objetos.
- Las características extrínsecas son aquellas que pueden considerarse externas a su definición como clase; por ejemplo, función, origen, localización o inventor de un objeto.

Un ejemplo es la característica "mineral" que se refiere a una propiedad intrínseca del concepto "polihalita", ya que la define como miembro de una clase y no necesita de otros conceptos para caracterizarla. No es el caso de la característica "se encuentra en Ische (Austria)" que describe la polihalita en relación con una situación, ajena a la clase de conceptos, que no afecta a su definición como miembro de los minerales (ibíd.).

En la norma UNE 1066 (AENOR, 1991: 2) las características extrínsecas se dividen en:

- Características de procedencia u origen (donde, a quién o como un objeto ha llegado a existir, a ser empleado o ser conocido)
- Características de destino o empleo (a qué se destina un objeto).

Felber y Picht (Felber y Picht, 1984) proponen una clasificación similar aparecida en la norma DIN 2330 (DIN, 1979: 6) y que prescinde de la primera división entre características intrínsecas y extrínsecas:

- Las características de estado, donde encontramos características de forma, de color, de posición o de tiempo.
- Las características relacionales, estas incluyen características de procedencia, de uso, de comparación, de valoración o de situación en el espacio.
- Las características funcionales, englobarán las características de potencia o rendimiento y las de empleo.

Cadiot y Nemo (Cadiot y Nemo, 1997a; Cadiot y Nemo, 1997b) introducen las nociones de *característica intrínseca* y *característica extrínseca* con el fin de explicar la polisemia desde la semántica léxica. Para los autores, las propiedades intrínsecas son estables y reflejan una lógica de clasificación, lo que permite la categorización de los nombres de los objetos en clases.

Las características extrínsecas, en cambio, surgen de relaciones que el locutor mantiene con el referente, por la forma en la que el locutor toma contacto con el objeto (Cadiot y Nemo, 1997a: 24).

Señalan que, dado que ningún objeto puede separarse de las relaciones que se mantienen con él, la descripción del objeto debe tener en cuenta de forma diferenciada las propiedades extrínsecas relativas a esas relaciones y las propiedades intrínsecas, que se suponen estables e independientes de éstas. La importancia de unas u otras varía según se trate de fenómenos naturales, artefactos u objetos abstractos. Ponen como ejemplo que en el fenómeno de la ‘noche’, entender la naturaleza física del mismo (y por tanto, sus propiedades intrínsecas) no impide que la palabra ‘noche’ pueda designar también el momento en que uno duerme, es decir, la relación (propiedad extrínseca) que uno mantiene con la realidad en cuestión, una parte del día establecida de forma arbitraria, una situación vista como oscura... Por eso, nadie creerá que es raro o metafórico que un explorador polar pueda decir ‘he dormido bien esta noche’ aunque nos haya explicado minutos antes que no se ha hecho de noche en cuatro meses (Cadiot y Nemo, 1997b: 128).

Tener en cuenta la diversidad de usos de las palabras pasa por la construcción del dominio de distribución de las propiedades extrínsecas y las propiedades intrínsecas, según dos ejes principales:

- el de la diversidad de relaciones posibles (propiedades extrínsecas) con un mismo ‘objeto’ (un solo conjunto especificado de propiedades intrínsecas)
- el de la diversidad de ‘objetos’ con los que se puede mantener una misma relación. Cadiot y Nemo (ibíd.) denominan unidad léxica al conjunto de usos de una palabra que se traducen por un mismo tipo de relación.

De acuerdo con Cadiot y Nemo, la diversidad de relaciones posibles con un objeto aparece cuando distinguimos entre la intensión y la extensión; lo que distingue al ‘Lucero del Alba’ del ‘Lucero Vespertino’ no es el objeto Venus con sus propiedades intrínsecas sino las relaciones de Venus (propiedades extrínsecas), el hecho de que se observe a Venus por la mañana o por la tarde.

Según estos autores, la semántica léxica debe, en su función de estudiar el sentido de las palabras, asumir que las propiedades intrínsecas son independientes de las propiedades extrínsecas. El sentido de las palabras se calcula a partir de las propiedades extrínsecas, ya que cuando se habla del sentido de las palabras en base a las propiedades intrínsecas no se habla de semántica. Lo que interesa a la semántica léxica no es aquello que responde a la pregunta “¿qué es x?”, que se respondería con las propiedades intrínsecas del objeto x, sino identificar cuáles son las propiedades extrínsecas que explican que un conjunto de objetos con propiedades intrínsecas diferentes se denominan de la misma manera. Dicho de otra forma, los autores señalan que la función de una palabra observada a través de sus usos, es la de establecer un acceso a los objetos a través de la relación que mantenemos con ellos y este acceso debe formularse en forma de propiedades extrínsecas.

Estas ideas presentadas por Cadiot y Nemo tienen como finalidad remarcar que el sentido de las palabras se debe explorar a través de sus propiedades extrínsecas. Para explicar la noción de propiedad extrínseca se basan en una serie de ejemplos de uso de *patates chaudes* (patata caliente).

- 1) *patates chaudes* désigne des pommes de terre à une certaine température (dans un rapport alimentaire)
- 2) *patates chaudes* désigne des pommes de terre à une temperature rendant difficile de s'en saisir ou douloureux de les garder en main ;
- 3) *patates chaudes* désigne dans certaines services sociaux des dossiers que l'on préfère ne pas ouvrir ou dont on s'empresse de se débarrasser – ou encore des personnes que l'on s'en renvoie d'un service à l'autre.

Los objetos designados en (1) y (2) tienen muchas propiedades intrínsecas en común pero el objeto designado en (3) no tiene propiedades intrínsecas en común con (1) y (2). Sin embargo, la relación que se mantiene con los objetos (1) y (2) es diferente mientras que la que se mantiene entre (2) y (3) es similar.

Con este y otros ejemplos los autores vienen a explicar que las propiedades extrínsecas son aquellas que identifican las relaciones que mantenemos con un objeto. Y con esto, entienden que la metonimia puede estudiarse desde las propiedades extrínsecas. Además, desde su punto de vista, no se puede estudiar un objeto sin tener en cuenta sus propiedades extrínsecas ya que las palabras están gracias al hombre y no tiene sentido estudiarlas sin atender a las relaciones que el hombre mantiene con los objetos que se denominan a través de ellas.

Gaudin (Gaudin, 2000) analiza lo mencionado por Cadiot y Nemo y explora qué tipo de características, intrínsecas o extrínsecas, ayudan a tratar mejor el concepto de 'polisemia'. Lo que viene a decir Gaudin es que las prácticas que tenemos con los objetos, las relaciones que mantenemos con ellos son más constantes que las categorías. Por ejemplo, un 'lecho' podrá dejar de ser un mueble pero siempre será un sitio donde tumbarse (Gaudin, 2000: 155).

Gaudin asegura que las propiedades intrínsecas están influenciadas por las clasificaciones jerárquicas, de hiperonimia, como estamos acostumbrados. Ciertas relaciones se establecen según la práctica y la cultura: un fresno es un tipo de árbol, la tristeza un tipo de sentimiento, la cuchara un tipo de cubierto. Las propiedades intrínsecas que los autores de las definiciones consideran aparecen por un sentimiento lingüístico espontáneo como primordiales. La consciencia lingüística no es apenas sensible a las virtualidades semánticas, al significado, se centra principalmente en la utilidad eficaz que constituye la relación de denominación. Según el autor, las terminologías están pensadas para funcionar como un conjunto de nombres-etiquetas de signos para los cuales el significado está encajado en la designación, y de las cuales ellas constituyen un terreno privilegiado para examinar el trasfondo de proposiciones como las que desarrollaron Cadiot y Nemo.

Desde su punto de vista, la polisemia se puede fundamentar en las propiedades extrínsecas de un objeto, que son las que permiten observar la relación de un objeto con la práctica. Sin embargo, advierten que las prácticas asociadas a los objetos, propiedades extrínsecas de éstos, no siempre son válidas para todos los locutores.

2.2.4 Características como categorías descriptivas

Bajo este epígrafe hemos reunido algunas de las tipologías que, desde nuestro punto de vista, agrupan las características a partir de otras características más generales.

Kageura en su trabajo sobre la formación de términos de la documentación clasifica un conjunto de morfemas japoneses en: entidades, cualidades, relaciones y actividades y, como hemos visto en §1.2.2, distingue las cualidades en etiquetas de características y valores de

características. Siguiendo esto, establece una clasificación de tipos de etiquetas y otra de tipos de valores.

En la clasificación de etiquetas de las características, Kageura propone distinguir entre aquellas que son propias del ámbito tratado, denominadas redefinidas, y aquellas que pertenecen al ámbito general, denominadas indefinidas. Además, distingue un tercer grupo de etiquetas que engloba las etiquetas de características de medida cualitativa o cuantitativa. Por ejemplo, la característica ‘condición’ se considera indefinida porque el concepto que representa en el ámbito de la documentación coincide con el del ámbito general. Por otro lado, ‘estructura’ aparece en el grupo de características redefinidas, porque alude a la propiedad concreta en que se distribuye u ordena una obra documental. En el grupo de características de medida cualitativa o cuantitativa encontramos ‘beneficio’, ya que este se puede medir, por ejemplo, en porcentajes.

Por otra parte, la clasificación de valores se basa en una subdivisión de categorías generales que agrupan los valores. Según Kageura (Kageura, 2002: 76), se trata de un gran grupo de categorías que a menudo son generales pero entre las que se pueden encontrar también categorías más específicas del ámbito de la documentación, como por ejemplo “modo representacional”. A continuación, presentamos el esquema de clasificación de valores y algunos ejemplos de valores de su trabajo (Kageura, 2002: 77):

Subgrupo de valores de Kageura	Ejemplos de Kageura
Spatial quality	Universal
Temporal quality	Current
Functional/situationalquality	Readable
Original situation	National
Shape	Triangular
Colour	Blue
Amount/quantity	One
Scale	Micro
Generality/abstraction	General
Negative	non-
Naturalness	Natural
Completeness/detailedness	Complete
Structure/mode	onelayer
Representationalmode	Analogue
Manner/means	one-by-one

Tabla 1. Tipos de valores de características (Kageura, 2002: 77)

En resumen, esta propuesta agrupa los nombres de características según la especialización de su concepto y los valores de las características a través de una clase o categoría más general.

La tipología utilizada en el trabajo de Vaupel (Vaupel, 2005) toma menos campos que la anterior. Su trabajo está orientado a la creación de un sistema de clasificación de conceptos del ámbito de la construcción y por tanto se centra en estudiar un patrón lógico, natural y común entre las propiedades de los elementos de la construcción que pueda utilizarse como base de un método sistemático y científico de clasificación de elementos de la construcción.

Clasificación de Vaupel
Propiedades materiales
Propiedades de función
Propiedades geométricas
Propiedades temporales

Tabla 2. Tipos de propiedades según Vaupel (Vaupel, 2005)

En su trabajo Vaupel señala que este modelo está inspirado en las cuatro formas básicas de concepción de la teoría de la percepción del filósofo Immanuel Kant. Las propiedades materiales se componen de características o clases de características que son productos físicos o

materiales, las propiedades de función recogen características relativas a la función de los conceptos materiales, las propiedades geométricas se componen de clases de formas o conceptos que describen espacios en tres dimensiones y finalmente, las propiedades temporales engloban las duraciones y puntos en el tiempo.

Esta clasificación está pensada para representar características concretas de un ámbito y por tanto, no resulta suficiente para representar otras características, por ejemplo, “calidad” o “posibilidad de suministro” ya que éstas se englobarían en grupos distintos a los expuestos arriba. Sin embargo, queremos destacar que esta clasificación a diferencia de la anterior, agrupa las características independientemente de que sean etiquetas de característica o valores.

Madsen y Thomsen (Madsen, 1998; Thomsen, 1998), por ejemplo, utilizan un criterio descriptivo en la tipología que sirve de punto de partida de su clasificación de características de dispositivos médicos. Esta clasificación agrupa las características a través de otras más generales como son:

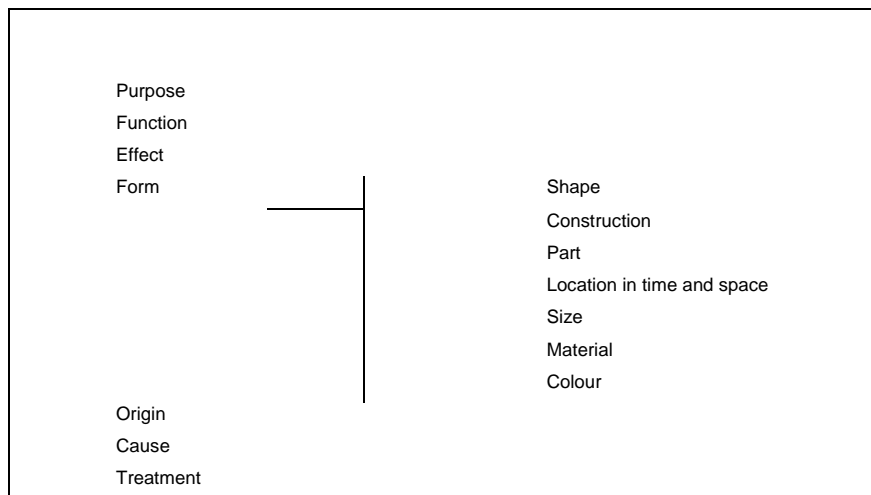


Tabla 3. Clasificación de Madsen (Madsen, 1998: 342)

Como vimos anteriormente, Madsen y Thomsen entienden que una característica es un par compuesto por un atributo y un valor y con esta base se plantean su clasificación con el fin de reflejar la relación entre el par y el concepto a definir (Thomsen, 1998: 356). En esta clasificación, además, se incluyen en un mismo sistema de representación las características de un concepto y las relaciones semánticas que existen entre un concepto y otro, ya que, las autoras consideran que una característica se puede representar como una relación semántica y viceversa.

Las autoras señalan que su tipología de característica describe los tipos de relación (representada por el atributo de la característica) que se establece entre el concepto a definir y el concepto asociado que sirve para definirlo (representada por el valor de la característica). En el ejemplo de Madsen sobre características del dominio de los seguros de vida, encontramos que la propiedad “coverage_period: lifetime”, forma parte del grupo de características de “forma” y dentro de este grupo es una característica de “situación en tiempo y espacio”. Así representan que el concepto “life insurance” (seguro de vida) y el valor, “lifetime” (vida) están relacionados mediante una característica de tipo “situación en tiempo y espacio” reflejada por el atributo “coverage period” (periodo de cobertura).

2.2.5 Características según el ámbito

Otra forma de clasificar las características de un concepto, sería agruparlas según el dominio al que pertenecen. Por ejemplo, la característica 'resistencia a la abrasión' formará parte del campo de la cerámica. El campo de especialidad donde se utiliza una característica proporciona una información adicional sobre su naturaleza y su concepto y resultará útil en su descripción.

Este criterio de clasificación se ha utilizado en terminografía para la descripción de conceptos. Así, cuando introducimos un concepto en una ficha terminológica añadimos el área de especialidad al que pertenece.

La Clasificación Decimal Universal es uno de los sistemas de clasificación por dominio más extendidos en el mundo. La Clasificación Decimal Universal es un sistema de clasificación del conocimiento que nace de la necesidad de ordenar y catalogar las obras en las bibliotecas. Esta clasificación se basa en la ordenación del conocimiento a través de dígitos. El conocimiento se enumera en 10 grandes grupos, y siguiendo el principio de subordinación lógica, cada uno de estos grupos puede subdividirse a su vez, y así sucesivamente descendiendo en el nivel jerárquico. Un ejemplo³ de tres niveles es el siguiente:

- 6 Ciencias Aplicadas. Medicina. Técnica.
- 66 Industria química. Tecnología química. Industrias afines.
- 666 Industria del vidrio, cerámica, cemento y hormigón.

Actualmente, es AENOR (Asociación Española de Normalización y Certificación) la responsable de la dirección y mantenimiento de la CDU en España y es miembro del Consorcio de la CDU.

La clasificación de la Biblioteca del Congreso⁴ (LCC, Library of Congress Classification) es otro ejemplo de clasificación por dominio y ha sido desarrollado por la Biblioteca del Congreso de los Estados Unidos. Aunque divide las materias en categorías amplias, la clasificación es esencialmente enumerativa en naturaleza, es decir, se basa en la acumulación de descriptores. Esta clasificación no pretende ser universal y en este aspecto ha recibido críticas por carecer de una base teórica sólida. Esta clasificación es fundamentalmente pragmática y las decisiones tomadas en su desarrollo estuvieron motivadas por las necesidades prácticas específicas de esa biblioteca más que por consideraciones epistemológicas. Esto hace que las ramas no estén desarrolladas con la misma profundidad unas que otras y resulte difícil aplicar en sistemas de organización de conocimiento. Un extracto de esta clasificación es el siguiente:

- T – Technology
- TP – Chemical technology
- TP785-869 Clay industries. Ceramics. Glass

2.2.6 Características equivalentes

En la norma UNE 1070, bajo el epígrafe «Tipos de caracteres» encontramos, además de las características esenciales, no esenciales, intrínsecas y extrínsecas, los caracteres equivalentes. Estas características se diferencian de las no equivalentes al observarlas desde el

³ Fuente del ejemplo: <http://www.mcu.es/libro/docs/TablaCDU.pdf>

⁴ La clasificación se puede consultar en <http://www.loc.gov/catdir/cpsol/lcco/>

punto de vista cognitivo. Se trata de características que, si bien desde un punto de vista lógico no se refieren al mismo concepto, desde el punto de vista ontológico sí conectarían con el mismo concepto:

«Caracteres equivalentes: caracteres diferentes que, no obstante, pueden sustituirse uno por otro en una comprensión dada sin que por ello resulte un cambio de extensión.» (AENOR, 1979: 6)

En la misma norma se pone de ejemplo la equivalencia entre 'equilátero' y 'equiángulo'. Un objeto equilátero es aquél cuyos lados son iguales, al mismo tiempo, un objeto equiángulo es aquél cuyos ángulos son iguales. Si analizamos estos dos conceptos, notamos que, en realidad, hacen el resultado de uno y otro da lugar al mismo tipo de objeto. De ahí que se considere que las características 'equilátero' y 'equiángulo' son equivalentes.

2.2.7 Recapitulación y observaciones

Hemos visto que una característica es un concepto que sirve para describir otros conceptos. Asimismo, hemos visto que las características se pueden clasificar de distintas formas que varían según el punto de vista o el criterio que se haya tomado en cuenta a la hora de clasificarlas. En este contexto, no todos los criterios funcionan del mismo modo, por ejemplo, hemos visto formas de clasificar la característica que requieren de otros conceptos además de la propia característica para ser interpretadas y otras en las que será suficiente la 'característica'. Aunque algunas de las clasificaciones que hemos presentado barajan distintos criterios de clasificación, sí podemos afirmar que cada uno de ellos está representando una dimensión del concepto 'característica'. Las distintas dimensiones de 'característica' reflejadas en las distintas formas de clasificación proporcionan información acerca de estas y servirán para describirlas desde el punto de vista terminológico. Considerando el concepto de multidimensionalidad explicado anteriormente, entendemos que estas distintas dimensiones de una característica aluden a distintas características de una característica.

Una característica puede ser esencial, esencial distintiva o no esencial, según su importancia en la descripción de un concepto y su importancia para distinguir un determinado concepto de otros en un mismo dominio. Resulta muy importante, como remarcan todos los autores, tener en cuenta que esta distinción requiere de un concepto referente para ser entendida y adecuada. Además, debemos tener en cuenta que una característica puede ser esencial cuando tiene un determinado valor o cuando sus valores se encuentran dentro de unos límites, esto nos hace reflexionar sobre qué hace esencial a una característica. Como demuestra Sager en la enumeración de características esenciales de 'mesa', ésta tiene unos límites dimensionales y de proporción, por lo que el valor de las dimensiones y proporción no es uno en concreto sino que puede ser uno entre unos límites. En cambio, otras veces, resulta esencial que la característica esté presente y además posea un valor concreto. Siguiendo con los ejemplos de Sager, la superficie de 'mesa' debe ser 'lisa'. En conclusión, podemos encontrarnos características que son esenciales porque tienen un valor en concreto y otras cuyo valor puede ser uno entre un rango.

Esto enlaza con el siguiente punto que distingue entre nombres de características y valores. Poco debemos añadir al respecto, esta distinción entre los dos elementos que componen una característica desde el punto de vista de su representación ha sido utilizada por numerosos autores.

La diferenciación entre características intrínsecas o extrínsecas resulta, desde nuestro punto de vista, difícil de explicar sin caer en la subjetividad. Notamos que las definiciones de característica intrínseca y característica extrínseca que hemos analizado son susceptibles de considerarse poco precisas y, en algunos casos, ambiguas cuando ponemos en común los criterios utilizados para diferenciar unas características de otras con otros aspectos como la multidimensionalidad. Definir intrínseco como algo que sirve para clasificar, no resulta efectivo si entendemos que un conjunto de conceptos se pueden clasificar según su ‘origen’, ejemplo de característica extrínseca por excelencia. Sin embargo, debemos reconocer que existe una tradición en las clasificaciones y categorizaciones que se basa en las características intrínsecas como criterio de categorización entre conceptos. El argumento se desprende de lo que señalan los distintos autores y es que las características extrínsecas aluden a conceptos que deben ser interpretados por el receptor para entender la relación que tiene con el concepto descrito, lo que dificulta que una categorización pueda ser entendida por todos los receptores.

Cadiot y Nemo denominan *propiedades extrínsecas* a aquellas que identifican las relaciones que mantenemos con un objeto. Por oposición, una propiedad intrínseca es aquella relativa únicamente al objeto y que no identifica ninguna relación entre el objeto y el mundo.

Además, Cadiot y Nemo señalan que son las propiedades extrínsecas las que ayudan a entender la metonimia. Desde un punto de vista terminológico, dos objetos con propiedades intrínsecas diferentes representan dos conceptos distintos aunque reciban la misma denominación. Dejando este hecho claro, sí podemos interesarnos por las propiedades extrínsecas de los conceptos con una intención similar con la que lo hacen Cadiot y Nemo: para estudiar si cuando dos conceptos se denominan de una misma forma se trata de un préstamo terminológico entre campos de especialidad, de una adaptación de un concepto general a un campo específico, etc. y basar nuestro estudio en el tipo de propiedad extrínseca que comparten estos dos conceptos.

Basándonos en la bibliografía consultada, queda probado que una característica no puede ser intrínseca y extrínseca al mismo tiempo. De la bibliografía también se desprende que el carácter intrínseco o extrínseco de la característica está vinculado a la característica como concepto y como elemento descriptor ya que una característica siempre será intrínseca o extrínseca independientemente del concepto al que describa pero hacemos uso de un contexto donde describe a un concepto a través de la característica para identificar su carácter intrínseco o extrínseco.

Otra rasgo de las características es que al ser conceptos pueden clasificarse mediante características más generales que ayudan a describirlas. Esto en realidad no es algo exclusivo de las características pero sí resulta de gran interés para nuestra trabajo de descripción. Ocurre lo mismo cuando consideramos que las características mantienen una relación con su ámbito, es decir, el contexto en el que se hace uso de una característica o campo de especialidad. Aunque las dos clasificaciones presentadas no tienen la finalidad de clasificar conceptos según el ámbito sino de clasificar obras, consideramos que se trata de clasificaciones que pueden servir para expresar el ámbito de una característica.

Finalmente, hemos introducido en este apartado el concepto de característica equivalente porque creemos que este tipo de trabajos pueden ayudar a detectar este tipo de características.

2.3 Relaciones de las características

Las características mantienen distintas relaciones con otros conceptos y también con otras características. En la bibliografía, podemos encontrar distintos autores que describen y clasifican las relaciones entre conceptos desde distintos enfoques como la terminología, la lingüística o la psicología (Lyons, 1977; Cruse, 1986; Winston et al., 1987; Nuopponen, 1994; Feliu, 2004; Maroto, 2007).

La definición que proponen Feliu y Cabré (Feliu y Cabré, 2002) concibe las relaciones como un nexo semántico entre dos o más unidades de conocimiento especializado. Esta definición se formaliza mediante la siguiente fórmula:

$$a R b, n$$

En esta fórmula, R representa a la relación y a y b son las clases conceptuales relacionadas, mientras que n refleja la recursividad que puede existir en algunas relaciones. Es decir, la posibilidad de que haya una relación R entre a y más elementos que se situarían al lado derecho de la fórmula (Feliu 2004: 30-34). Esta fórmula está basada en Otman (Otman, 1996: 55). Según Otman, una relación permite establecer vínculos nocionales entre varios conceptos y se designan mediante un nombre, por ejemplo, tipo-de.

2.3.1 Propiedades matemáticas de las relaciones conceptuales

Las relaciones también se pueden observar como objetos matemáticos formados por pares (o series) de conjuntos que representan conceptos o palabras (Evans, 1988: 24). Al ser observadas de este modo, los autores se fijan en sus propiedades matemáticas con el fin de describirlas. En los trabajos de distintos autores (Lyons, 1977; Cruse, 1986; Winston et al., 1987; Nuopponen, 1994; Feliu, 2004; Maroto, 2007) podemos encontrar menciones a las siguientes propiedades matemáticas:

- *Transitividad.* Suponiendo que exista una relación R que vincula a un concepto a con un concepto b y al mismo tiempo existe la misma relación R entre b y c . Se considera que la relación R es transitiva, si ello implica que a y c están relacionados mediante la misma relación R . Por ejemplo, si a es más pequeño que b y b es más pequeño que c , y la relación es transitiva, entonces a es más pequeño que c .
- *Simetría.* Si una relación R es simétrica, cuando un concepto a tiene una relación R con b , el concepto b tiene la misma relación R con a . Por ejemplo, si a está casado con b , b está casado con a .
- *Recursividad.* Si una relación R es recursiva, un concepto a tiene una relación R con más de un elemento ($b, c, d...$).
- *Reflexividad.* Una relación es reflexiva si se cumple que a tiene una relación R con el mismo concepto a , por ejemplo, 'Álex' tiene la misma edad que 'Juan' (Juan tiene la misma edad que Juan, Álex tiene la misma edad que Álex).
- *Inversa.* Cuando una relación $R1$ entre un concepto a y un concepto b es la opuesta a una relación $R2$ entre el concepto b y el concepto a , $R2$ se denomina *relación inversa* de $R1$. Por ejemplo, si a es sobrino de b y b es tío de a , podemos decir que *sobrino-de* es una relación inversa a *tío-de*.

Además de estudiar las propiedades de las relaciones, distintos autores han estudiado sus similitudes y diferencias con el fin de clasificarlas y ordenarlas mediante tipologías. Feliu realiza un trabajo exhaustivo sobre los distintos tipos de relaciones conceptuales que tiene como resultado un catálogo en el que se describe cada relación, incorporando información sobre sus propiedades matemáticas, las clases conceptuales implicadas, el número de argumentos oraciones y ejemplos en catalán (Feliu, 2004: 30-34). Maroto (Maroto, 2007) se basará en este catálogo y en otros trabajos para formalizar un gran número de relaciones conceptuales en una ontología. A continuación, vamos a resumir el catálogo de relaciones que recogió Maroto (Maroto, 2007: 315), haciendo hincapié en aquellas relaciones en las que participa una característica.

2.3.2 Relaciones lógicas

Las relaciones lógicas permiten establecer jerarquías o taxonomías entre conceptos en virtud de las características que comparten y que los diferencian. Dentro de las relaciones lógicas, se distingue la relación que se establece entre un superordinado y sus subordinados, vista como una relación vertical, y la que vincula a subordinados entre ellos, vista como una relación horizontal (Sager, 1993: 55-63).

2.3.2.1 Relación superordinado-subordinado

Es la relación que se establece entre un concepto *a* que posee un grado de abstracción superior a un concepto *b*. El concepto *b* posee las mismas características que el concepto *a* más alguna adicional que aumenta su intensión y lo hace más específico.

Esta relación se puede establecer entre entidades, características y actividades siempre que los conceptos situados a ambos lados de la relación pertenezcan a la misma clase. Un ejemplo relativo a las propiedades sería *coeficiente de fricción* – *coeficiente de fricción dinámico* (Maroto, 2007: 344).

2.3.2.2 Relación subordinado-subordinado

En este caso se trata de conceptos que comparten características con el superordinado común y por lo tanto, se agrupan bajo un mismo superordinado pero tienen al menos un rasgo que los diferencia entre sí, y que los diferencia a su vez del superordinado. Cuando dos o más conceptos cumplen este criterio se les denomina *co-hipónimos*.

Las relaciones de tipo subordinado-subordinado son relaciones son transitivas y simétricas. Además, pueden ser recursivas, es decir un concepto *a* puede tener una relación *subordinado-subordinado* con más de un concepto (*a R b, c, d...*). Igual que con la relación anterior, esta relación se puede establecer entre entidades, propiedades y actividades siempre que los conceptos situados a ambos lados de la relación pertenezcan a la misma clase. Un ejemplo relativo a las propiedades sería *curvatura* – *grosor* (ibídem: 344).

2.3.3 Relaciones parte-todo

Hacen referencia a la relación que se establece entre un todo y sus partes, por esto, también encontramos otras denominaciones para esta relación como relaciones meronímicas o relaciones partitivas. En este aspecto, Feliu señala este tipo de relaciones parte-todo se dan entre distintos tipos de concepto, pero que siempre que el concepto implicado sea de tipo característica deberá mantener una relación con uno de tipo entidad (Feliu, 2004: 38).

2.3.3.1 Relación componente funcional-objeto

Se establece entre los componentes que desempeñan una función dentro del todo del que forman parte. Se distingue entre componentes canónicos, cuando son necesarios para la constitución del todo, y facultativos, cuando no son imprescindibles para que el todo siga siendo el todo. Esta relación suele ser recursiva, puede ser transitiva o no pero nunca será simétrica. Generalmente, en esta relación tanto las partes como el todo son ejemplificados mediante conceptos de tipo entidad. Un ejemplo de Maroto es *arista-baldosa* (ibídem: 344).

2.3.3.2 **Relación miembro-colección**

Hace referencia a la relación que se establece entre partes que están próximas en el espacio o tienen una conexión social con respecto al todo. Nuevamente, este tipo de relación se establece entre conceptos de tipo entidad y es su similitud con el resto de los miembros lo que determina su pertenencia a la colección. En cuanto a sus propiedades matemáticas esta relación no será transitiva ni simétrica. Un ejemplo de Maroto (ibídem: 350) es *tesela-mosaico*.

2.3.3.3 **Relación porción-masa**

Es la relación entre una parte y un todo que son homeoméricos, es decir, comparten la misma naturaleza. Los conceptos vinculados mediante esta relación suelen ser entidades, como por ejemplo *porción-pizza* (Winston y otros: 1987: 421). Esta relación es transitiva y no simétrica.

2.3.3.4 **Relación material-objeto**

Indica la relación entre una parte intrínseca y el todo al que pertenece. Esta parte no puede separarse del todo ya que es consustancial al mismo, precisamente este rasgo de no ser separable es el que distingue a esta relación de otras relaciones parte-todo. Como otras relaciones similares, solo existen ejemplos de que ocurra entre entidades. Respecto a sus propiedades matemáticas, se trata de una relación transitiva, no simétrica y un ejemplo de Maroto (ibídem: 352) sería *arcilla-baldosa cerámica*.

2.3.3.5 **Relación etapa-proceso**

Sirve para denominar a la relación que se establece entre las fases de una actividad y la actividad entendida como un todo. Las clases conceptuales relacionadas suelen ser actividades. En el trabajo de Maroto, encontramos el ejemplo *ensayo-control de calidad* (ibídem: 353). Estas relaciones son transitivas y no simétricas.

2.3.3.6 **Relación característica-actividad**

Se denomina así a la relación que se establece entre las características de una actividad y el proceso en el que se inscriben. En esta relación una característica *a* mantiene una relación con una actividad *b*, de modo que no existe simetría pero sí puede existir recursividad o transitividad (ibídem: 353). Un ejemplo típico de esta relación es la relación entre *pagar* y *comprar*, pagar es una parte de la actividad comprar (Winston et al., 1987: 421).

2.3.3.7 **Relación lugar-área**

Señala la relación que establece entre un área espacial y localizaciones concretas dentro de esa área. Podemos distinguir esta relación porque el lugar y el área compartirán la misma naturaleza. Las clases conceptuales implicadas suelen ser entidades, por ejemplo, *Atzeneta-Castellón*. Esta relación es transitiva y no simétrica.

2.3.3.8 **Relación parte-parte**

Es una relación horizontal que se establece entre las partes que constituyen un todo. En este caso, la relación es más abstracta y agrupa todas las relaciones que pueden establecerse de

tipo *componente-componente*, *miembro-miembro*, *porción-porción*, *material-material*, *etapa-etapa* y *característica-característica* con respecto a su todo.

En este caso, dependiendo del tipo de parte-parte, las clases conceptuales que participan de la relación serán entidades, actividades o propiedades. Esta relación es transitiva y simétrica y puede ser recursiva (Maroto, 2007: 355). Un ejemplo de relación parte-parte entre entidades sería: *cara vista-arista*.

2.3.4 Relaciones secuenciales

Maroto continúa su clasificación adentrándose en las relaciones secuenciales son aquellas que se establecen por localización o sucesión en el espacio o el tiempo de los elementos que vincula (ibídem: 356-360).

2.3.4.1 *Relación concepto-lugar*

Determina la relación que se establece entre dos conceptos que se ubican juntos en el espacio. En esta relación pueden participar conceptos de tipo entidad o actividad. Un ejemplo sería la relación entre *baldosa cerámica* y *suelo*.

2.3.4.2 *Relación entre concepto-lugar al que se dirige*

Sirve para denominar la relación entre dos conceptos donde se indica la dirección en la que se produce una actividad (Feliu, 2004: 43). Las clases conceptuales implicadas suelen ser entidades y actividades, un ejemplo sería la relación entre un vector y la dirección hacia la que apunta. Estas relaciones pueden ser transitivas y simétricas.

2.3.4.3 *Relación concepto-concepto simultáneo*

Es la relación entre dos actividades o entidades que se suceden en el mismo espacio de tiempo. Esta relación es simétrica y transitiva. Un ejemplo sería la relación entre el momento en que aparece una teoría y la teoría en sí misma.

2.3.4.4 *Relación concepto anterior-concepto posterior*

Indica la relación entre dos conceptos, actividades o entidades, que se suceden en el tiempo de forma consecutiva. Esta relación es no simétrica y transitiva, un ejemplo sería la relación entre *anestesia* y *operación*.

2.3.5 Relaciones argumentales y circunstanciales

Agrupan las relaciones paradigmáticas que se establecen entre los predicados y sus argumentos así como las circunstancias en las que se produce un predicado.

2.3.5.1 *Relación proceso-agente*

Es la relación que se establece entre un proceso (actividad) y la entidad o entidades que llevan a cabo dicho proceso. Esta relación es no transitiva y no simétrica. Por ejemplo, la relación entre *alicatar* – *alicatador*.

2.3.5.2 *Relación proceso-producto*

Sirve para denominar la relación que se establece entre el proceso (actividad) que se lleva a cabo para obtener un producto (entidad). Esta relación no es simétrica y puede ser transitiva o no. Un ejemplo sería la relación que se establece entre *moldeo* y *baldosa cerámica*.

2.3.5.3 *Relación proceso-paciente*

Determina la relación que se establece entre un proceso (actividad) y la entidad sobre la que se lleva a cabo dicho proceso. Estas relaciones no son transitivas ni simétricas. Un ejemplo es la relación entre *colocación* y *baldosa cerámica*.

2.3.5.4 **Relación proceso-instrumento**

Esta relación se establece entre un proceso (actividad) y el instrumento (entidad) empleado para llevar a cabo dicho proceso. Esta relación es no transitiva y no simétrica y un ejemplo de Maroto (ibídem: 364) es colocación en *capa gruesa – cemento aditivado*.

2.3.5.5 **Relación proceso-estado**

Expresa la relación entre un proceso (actividad) y el estado final que alcanza el paciente de dicho proceso, que puede ser una propiedad o entidad. Esta relación es no transitiva y no simétrica y su relación inversa sería estado-proceso. Un ejemplo que nos viene a la mente sería la relación entre *extrusión* y *baldosa extruida*.

2.3.5.6 **Relación causa-efecto**

Es la relación que se establece entre una causa y el efecto que produce. Las clases conceptuales entre las que se puede establecer este tipo de relación son entidades, actividades y propiedades y no es necesario se establezca entre miembros de un mismo tipo. Esta relación no puede ser simétrica pero sí puede ser recursiva o transitiva. Un ejemplo con características es *porosidad abierta – absorción de agua*.

2.3.5.7 **Relación objeto – uso**

Es la relación que se establece entre un objeto y su uso. En estas relaciones que no pueden ser simétricas ni transitivas, encontramos entidades, actividades y propiedades relacionadas entre sí o entre ellas. Ejemplos de Maroto son *barniz y uso decorativo, baldosa-revestimiento de suelos o capa de desolidarización – aislamiento del soporte*.

2.3.6 **Otras relaciones**

En este apartado, Maroto (ibídem: 319) incluye las relaciones que no encajan dentro de los cuatro grupos anteriores pero se establecen entre conceptos y pueden resultar útiles para estructurar conceptos de un campo de especialidad de la cerámica. A continuación presentamos las relaciones que propone con sus ejemplos.

2.3.6.1 **Relación fenómeno – medida**

Indica la relación que se establece entre un fenómeno o característica y la unidad de medida empleada para expresar dicho fenómeno. En esta relación los conceptos relacionados pueden ser actividades, entidades o propiedades y normalmente, no se establece entre miembros de un mismo tipo. Esta relación no puede ser simétrica ni transitiva pero sí puede ser recursiva. Un ejemplo de relación de tipo fenómeno-medida sería *carga de rotura – Newton*.

2.3.6.2 **Relación objeto-característica**

Esta relación es la que vincula a una entidad con una característica que define a dicha entidad. Esta relación se establece entre entidades y características. Esta relación es la inversa de la relación característica-objeto, puede ser transitiva pero no simétrica. Por ejemplo, *azulejo-resistencia a la helada*.

2.3.6.3 **Relación asociativa**

Es una relación más abstracta que las anteriores y describe la relación que existe entre dos conceptos y cuya naturaleza no se corresponde con ninguno de los tipos de relación

identificados en la bibliografía. Puede establecerse esta relación entre cualquier tipo de concepto y en cualquier combinación y se basa en la asociación entre dos conceptos en un determinado contexto, puede ser simétrica o no al igual que puede ser transitiva o no. Un ejemplo sería *embaldosado directo - capa de desolidariación*.

2.3.6.4 **Relación de sinonimia**

Es la relación que indica que existe una equivalencia conceptual entre dos o más denominaciones que refieren a un mismo concepto. Todos los tipos de conceptos pueden tener esta relación que es simétrica y transitiva. Además, la sinonimia puede ser recursiva y es la relación inversa de sí misma. Ejemplos son, *cuerpo-soporte, capa aislante-capa de aislamiento, canto-arista*.

2.3.7 **Recapitulación y observaciones**

En este apartado hemos recogido aquellas relaciones que permiten describir la naturaleza de una característica en interacción con otros conceptos. Como hemos visto, existen un número considerable de relaciones en las que participa una característica. También observamos que las características mantienen relaciones con los demás tipos de conceptos además de con otras de su mismo tipo.

En este aspecto, insistimos en resaltar que desde un punto de vista de la representación de conceptos, la distinción entre relación y característica es difusa (§2.1.4 y 2.1.5). Basándonos en esta idea, podemos considerar por ejemplo que la relación proceso-paciente, se puede representar en la descripción de un proceso en forma de característica a través del atributo “tiene paciente”, asignado como valor del atributo un paciente.

3 Representación de conceptos en terminología

Entendemos por representación de conceptos, el proceso de expresar mediante signos lingüísticos o símbolos el significado de un término. Tal y como lo resume Monterde (Monterde, 2002: 32) la representación de conceptos es la expresión lingüística o no lingüística, o una combinación de ambas, que permite delimitar un concepto entendido como una unidad de conocimiento con una combinación de características única.

A continuación, analizamos distintos tipos de representación de conceptos utilizados en Terminología donde veremos distintas formas de manifestar las características del concepto en el campo de la representación conceptual.

3.1 Las definiciones terminológicas

Cabré (Cabré, 1992: 193) explica que un concepto especializado se puede representar, por ejemplo, en diccionarios a través de una definición o una ilustración.

La ilustración es una unidad icónica que reproduce la idea que se tiene de una determinada clase de objetos de la realidad. La definición, en palabras de Cabré, es una fórmula lingüística que se propone describir el concepto que representa una denominación.

La norma ISO 704 (ISO, 2000) señala que una definición es «una descripción completa, normalmente a través del lenguaje, de una noción que utiliza otras nociones ya conocidas». La UNE-ISO 1087 (AENOR, 2009) además añade que una definición es «un enunciado que describe una noción y que, dentro de un sistema nocional, permite diferenciarla de otras nociones».

En este punto, Cabré (Cabré, 1992: 195) señala que es importante distinguir entre definiciones de tipo lingüístico, de tipo ontológico o enciclopédico y de tipo terminológico. Estos tres tipos se distinguirán por el objeto que describen y por los contenidos que expresan.

En lo relativo al objeto que describen, la autora explica que cada tipo de definición se refiere a una misma realidad vista desde perspectivas distintas. El objeto de la definición de tipo lingüístico se centra en el signo lingüístico; la definición de tipo ontológico se centra en el objeto de la realidad y la definición de tipo terminológico se centra en el concepto del sistema nocional de un campo de especialidad.

En lo relativo a los contenidos que expresan, la autora afirma que los tres tipos recogen distintas características. En una definición lingüística, no se suelen incluir todas las características de una noción, como es el caso de la terminológica, sino solamente aquellas que resultan relevantes para distinguirla de otras nociones. El objetivo de la definición lingüística será distinguir entre nociones. La definición ontológica incluye todos los aspectos peculiares de una noción, intrínsecos y extrínsecos, esenciales y complementarios, sin importar su relevancia para definirla como clase. Una definición terminológica, es descriptiva y no se hace por oposición, describe la noción en referencia exclusiva a un dominio de especialidad y no en referencia a un sistema lingüístico.

A pesar de esto, Cabré admite que la distinción es más nítida en la teoría que en la práctica, ya que podemos encontrar definiciones a medio camino entre estos tipos en algunas obras de especialidad.

De acuerdo con las recomendaciones del TERMCAT (TERMCAT, 2009) formalmente, la definición terminológica se compone de una única frase. En esta frase debemos incluir los rasgos esenciales de término a definir, se trata de poder caracterizar el término con respecto a otros términos similares del mismo campo conceptual, utilizando la nomenclatura del propio campo. En la definición incluiremos un definidor que actúe como descriptor (con la misma categoría sintáctica) antes de incluir los rasgos característicos. Los rasgos característicos, que modificarán al definidor, deben seguir un orden semántico; primero, se incluirán los rasgos intrínsecos y a continuación, los extrínsecos. En el caso de los rasgos extrínsecos, desde el TERMCAT se recomienda que las definiciones de los adjetivos incluyan información sobre su antecedente o sujeto (ibídem: 105). Por ejemplo, en la definición de fastigiado se incluye información sobre el sujeto (árboles) al final de la definición (ibídem: 106):

Fastigiado –da Adj. Que tiene las ramas rectas y cercanas al tronco de manera que la copa es larga y estrecha, dicho de los árboles.

A la hora de redactar la definición, el TERMCAT recomienda evitar el uso del término a definir dentro de la misma definición. Una forma de elaborar la definición es hacer uso de paráfrasis que representen la equivalencia semántica y sintáctica. Para comprobar si una definición es adecuada se puede recurrir a lo que denominan prueba de sustitución y consiste en comprobar si la definición puede sustituir a la denominación que define en un contexto de uso cualquiera (ibídem: 211).

Desde el TERMCAT aceptan la existencia de casos donde resulta interesante incluir información enciclopédica en las fichas de términos pero recomiendan que esta información aparezca en la definición de forma diferenciada y detrás de la definición terminológica. El término hará referencia a un concepto único y por tanto, es recomendable utilizar un único descriptor para iniciar la frase. En el caso de que haya sinónimos, éstos aparecerán en la misma ficha como denominaciones de un mismo término y con una única definición para todos.

Entre los aspectos que debemos evitar en las definiciones se encuentra la circularidad y la redundancia. En el caso de la circularidad, dos o más términos no podrán contener las denominaciones de otros términos de forma que ninguno quede definido de forma independiente con respecto a los otros. En cuanto a la redundancia, el TERMCAT recomienda evitar en una definición el uso de rasgos que ya constan en la definición del definidor correspondiente.

3.2 Las clasificaciones jerárquicas

El trabajo de Cabré (Cabré, 1992: 227) también explica que las clasificaciones conceptuales jerárquicas se basan en la estructuración vertical de los conceptos en clases y subclases. La autora señala que los conceptos se ordenan de general a particular y suelen representarse en una estructura arbórea, donde cada nudo del árbol, que representa a una clase conceptual, se suele codificar alfanuméricamente. Según Cabré, la clasificación más utilizada parte de la Clasificación Decimal Universal (AENOR, 2000) que adopta un esquema decimal con el cual pretende representar todos los campos del saber y se actualiza permanentemente por un comité del Instituto Internacional de Documentación (FID). A continuación, mostramos un fragmento como ejemplo de esta clasificación:

.1 Enfermedades del sistema circulatorio. Hematología

.11 Pericardio

- .12 Corazón
- .13 Arterias
- .14 Venas
- .15 Sangre. Plasma. Sérum
- .155 Anemias. Leucemias
- .16 Vasos capilares
- .2 Enfermedades del sistema respiratorio
- .21 Otorrinolaringología. Rinología
- .22 Laringología. Laringe
- .23 Tráquea y bronquios
- .24 Pulmones
- .25 Pleura
- .26 Diafragma

Según Cabré, las clasificaciones jerárquicas tienen la ventaja de representar un cuadro general lógico de la estructuración conceptual de las materias, fácil de utilizar para, por ejemplo, indexar documentos. Pero señala que son claramente insuficientes para explicar la multidimensionalidad de las relaciones conceptuales y por tanto, resultan redundantes.

3.3 Los tesauros

Siguiendo a Cabré (Cabré, 1992: 230), los tesauros son documentos constituidos por listas de términos normalizados, relacionados entre sí semánticamente, que representan las relaciones entre los conceptos que representan los términos. En un tesoro, un determinado campo del saber recibe una determinada organización semántica. De acuerdo con la autora, un tesoro se caracteriza por tres aspectos: sus contenidos, su estructura y su función.

Desde el punto de vista de contenido, un tesoro contiene una serie de términos relacionados entre sí semánticamente en tres dimensiones: la sinonimia, la hiponimia y la asociación. Cabré explica cómo la normalización de los términos que constituyen los tesauros, denominados descriptores, comporta la priorización de una denominación sobre las demás alternativas, el control de la sinonimia y la regularización de la forma de representación de los descriptores (masculino, singular, forma sustantiva).

Desde el punto de vista de la estructura, la autora explica que un tesoro es un documento altamente formalizado que consta de un vocabulario controlado y dinámico y de una serie de relaciones conceptuales expresadas formalmente a través de expresiones codificadas.

Desde el punto de vista de la función, la autora señala que los tesauros son instrumentos de control que regulan el uso del lenguaje natural para facilitar la recuperación de información.

Un tesoro consta básicamente de lo que Cabré denomina “artículos” integrados por tres tipos de elementos:

- una entrada, que es un término

- un conjunto de descriptores, que también son términos
- un conjunto de relaciones entre la entrada y cada descriptor

Las relaciones pueden ser de equivalencia, sinónimos y cuasi-sinónimos como por ejemplo: *EP* (entrada principal) mamífero, *UP* (utilizar por) Mammalia. Pueden ser de subordinación, por ejemplo: *EP* (entrada principal) Física, *TE* (término específico) física nuclear. O finalmente, las relaciones pueden ser de asociación, relacionando dos términos de manera general, por ejemplo: *EP* (entrada principal) cloroplastos, *TR* (término relacionado) fotosíntesis.

Finalmente, Cabré apunta que un tesoro no tiene una forma única, sino distintas formas que además no son excluyentes entre sí sino que resultan complementarias unas con otras. Por ejemplo, podemos encontrar un tesoro con forma de un diccionario conceptual que contenga un listado de términos y sus descriptores vinculados por relaciones semánticas, también podemos encontrar tesoros ordenados alfabéticamente o tesoros ordenados temáticamente.

3.4 Los bancos de datos terminológicos

Son herramientas muy útiles para los terminólogos tanto por para las búsquedas generales y especializadas de contenidos como para la realización de trabajos aplicados a cualquiera de las ramas del lenguaje y a la construcción de sistemas expertos (Cabré, 1992: 365)

El concepto de banco de datos terminológico ha ido cambiando con el tiempo, gracias a la evolución de las tecnologías informáticas y de la sociedad con respecto al uso del ordenador. Asimismo, ha habido una especialización en todos los campos con respecto a las tecnologías informáticas.

Un banco de datos es un conjunto estructurado de informaciones que están organizadas en registros. Cada registro consta de múltiples informaciones estructuradas en campos. Su creación requiere un proceso sistemático que se inicia con la concepción del modelo de banco y termina con la implementación de éste en un sistema informático. La elaboración de un banco de datos está determinada principalmente por el tipo de información que va a contener, las necesidades y los usuarios a los que va dirigido; por eso, encontramos diferencias entre los bancos de datos bibliográficos o documentales, los destinados a ser procesados por sistemas expertos, etc.

Los bancos de datos terminológicos, concretamente, se componen de un conjunto de información estructurado y organizado sobre las unidades de significado y designación de una área especializada y tiene como finalidad responder a las necesidades de un grupo definido de usuarios, por ejemplo, traductores especializados. Suelen constar de una base de datos principal, que contiene los términos del área y de algunas bases de datos complementarias que contienen información que complementa a alguno de los términos y mantienen algún tipo de relación con la principal y a veces, entre ellas.

La información terminológica que contienen suele proceder de publicaciones lexicográficas y terminológicas (diccionarios, léxicos, enciclopedias...) y la función de estos bancos suele ser facilitar a los traductores una herramienta de consulta única que integra muchas fuentes, fácil de consultar y fiable. Tradicionalmente, los bancos de datos utilizan las bases de datos como software que a su vez han ido evolucionando hasta permitir la consulta de bancos de datos on-line o con posibilidades de búsqueda más complejas.

3.5 Bases de conocimiento terminológico y sistemas orientados a ontologías

Condamines (Condamines, 2003: 46) señala que las bases de conocimiento terminológico forman parte de una visión interdisciplinar de la terminología. Esta denominación toma fuerza a través de los trabajos de Meyer et al. (Meyer et al., 1992a) pero existen menciones anteriores sobre esta idea en trabajos anteriores donde se vincula la terminología a la ingeniería artificial. Felber (Felber, 1987: 91), por ejemplo, señaló que dado que la inteligencia artificial y los sistemas expertos están dirigidos al tratamiento de los sistemas de nociones, sería recomendable un acercamiento entre la teoría general de la terminología y la informática. Esta voluntad de acercarse también se puede observar, según Condamines (ibíd.: 46), desde la ingeniería donde los informáticos empiezan a darse cuenta de la necesidad de apoyarse en los textos para «extraer y modelar» el conocimiento.

Para definir el concepto de *base de conocimiento terminológico* (BCT), Condamines establece una primera distinción entre ésta y una *base de datos terminológica* (BDT). Para ello, remarca tres aspectos que caracterizan una base de conocimiento terminológico. Por una parte, la representación en red de los conceptos, en una base de conocimiento terminológico, se realiza mediante una representación racional de los conceptos que en el caso de las bases de datos terminológicas solo se ve reflejada en forma de definición. En una BCT, las redes nocionales cobran un protagonismo que no existe en el caso de una BDT.

En segundo lugar, Condamines apunta hacia el significado de concepto en Terminología con respecto a la Ingeniería Artificial u otras disciplinas. En una BCT, tal y como la entiende Condamines, el concepto se distingue del término y es través de la representación de las relaciones de un concepto con otro como se va elaborando su descripción. El papel del término será representar el signo lingüístico con que se denomina un concepto, de forma que se explique la polisemia o sinonimia.

Un tercer aspecto que caracteriza una BCT, según la autora, es que este tipo de sistemas no necesariamente van a tener en cuenta el uso de los términos. Al representar los conceptos mediante redes nocionales y distinguir entre término y concepto no existe una necesidad de recurrir a los textos para recoger el conocimiento, ya que este puede proceder de los mismos expertos de un ámbito. No obstante, proyectos de este tipo, como Cogniterm, sí tienen en cuenta los textos como soporte al conocimiento representado. En este tipo de proyectos, se intenta establecer un vínculo entre la interfaz de acceso al conocimiento terminológico y los textos que hacen referencia a estos conceptos.

Las bases de conocimiento terminológico se consideran muy cercanas a lo que se entiende en Ingeniería del Conocimiento como ontologías. Autores con un perfil terminológico, definen ontología como: «a knowledge repository in which categories (terms) are defined as well as relationships between these categories» (Temmerman y Kerremans, 2003: 3). Y añaden que una ontología se compone de un conjunto de conceptos, axiomas y relaciones que describen un dominio de interés. Además, se explica que las ontologías se pueden representar mediante lenguajes naturales, que denomina *ontologías lingüísticas*, o mediante ontologías formales.

Condamines (ibídem: 56) hace referencia a las diferencias entre formalización desde un punto de vista de la Ingeniería del Conocimiento y desde un punto de vista lexicológico para distinguir una ontología, tal y como se describe desde la Ingeniería, de una base de conocimiento terminológico. Desde un punto de vista de la dimensión formal, las

modelizaciones de conceptos en forma de ontologías no difieren en gran medida de las planteadas desde un enfoque terminológico, ya que en ingeniería se hace uso de lenguajes inspirados en las redes semánticas como los grafos conceptuales o la lógica proposicional.

El vínculo de las ontologías al dominio descrito es otro aspecto en común con las BCT. Las ontologías generales han demostrado ser poco realistas y la tendencia actual es la de elaborar ontologías de dominio enfocadas a una aplicación concreta. Esto conlleva, sin embargo, que estas ontologías sean poco reutilizables. Para solucionar este problema, Condamines (ibídem: 57), propone utilizar el análisis textual como recurso para mejorar el proceso de creación de ontologías, de forma que la capitalización de conocimiento se enfoque desde los métodos para elaborar modelos de ontologías.

Madsen et al. (Madsen et al., 2010) señalan que las ontologías terminológicas difieren de otros tipos de ontología porque comprenden especificaciones de sus rasgos y criterios de subdivisión distintos y consideran el término *ontología terminológica* como sinónimo de *sistema de conceptos*. Desde su punto de vista, las ontologías sirven para clarificar conceptos. Según Madsen et al., en lugar de usar términos como *hiperónimo*, *hipónimo* y *co-hipónimo*, en las ontologías terminológicas estos conceptos se denominan *concepto subordinado*, *concepto superordinado* y *concepto coordinado*. En las ontologías, las características de los conceptos se presentan en forma de pares de atributo y valor, y ponen como ejemplo, [TARGET GROUP: population]. En base a estas especificaciones, se introducen criterios de subdivisión que permiten ilustrar las diferencias entre un conjunto de conceptos coordinados.

A continuación, vamos a presentar algunos ejemplos de trabajos que han desarrollado sistemas orientados a la representación de conocimiento desde un contexto de Terminología.

3.5.1 Cogniterm

Cogniterm es un proyecto desarrollado en la Universidad de Ottawa por Meyer et al. (Meyer y otros, 1992a; Meyer y otros, 1992b) que tiene la finalidad de construir un banco de conocimiento terminológico.

El banco de conocimiento terminológico (BCT) visto tal y como se propone en Cogniterm se enfoca dentro de la dimensión cognitiva de la terminología (Meyer et al., 1992a: 956-960). Esencialmente, se define como un híbrido entre una base de datos terminológica convencional y una base de conocimiento.

Cada concepto está representado en una estructura de tipo marco denominada *descriptor conceptual* (CD) que tiene dos grandes categorías de información. La categoría de *información conceptual* es el componente de la base de conocimiento, donde se enumeran las categorías conceptuales y sus valores. Cada CD se estructura normalmente mediante jerarquías basadas en la herencia. La categoría *información lingüística* es el componente donde se incluye información estrictamente lingüística, como la que podemos encontrar normalmente en una base de datos terminológica convencional.

Desgraciadamente, en este trabajo no hemos podido acceder al sistema desarrollado por Cogniterm y nos basamos en la bibliografía existente al respecto (Meyer et al., 1992a; Meyer et al., 1992b; Skuce y Lethbridge, 1995). Gráficamente, la BCT se visualizaría mediante distintas redes semánticas, donde se pueden mostrar tanto las relaciones jerárquicas (genérico-específico, parte-todo) como las no jerárquicas. Como la adquisición de conocimiento normalmente se basa en un subdominio cada vez, las sub-ventanas pueden mostrar únicamente una parte restringida

de la estructura de conocimiento (sub-árbol). Este modo permite también mostrar conceptos de acuerdo a una determinada 'dimensión' de la realidad.

Como ayuda a la construcción de definiciones y especialmente para ayudar a determinar las características diferenciadores, CODE4 ofrece una Matriz de comparación de características que presenta la unión de todas las características de conceptos coordinados (conceptos que comparten un mismo padre en una jerarquía) con la exclusión de aquellas que son idénticas en todos los coordinados.

Finalmente, la navegación a través de Cogniterm se realiza a través del navegador de CODE4 que permite acceder al conocimiento bien a través de los nombres de los conceptos o de los nombres de las características, ambos se pueden observar ordenados por concepto (jerárquicamente) o alfabéticamente.

De acuerdo con esta propuesta, las diferencias entre una BDT y una BCT se pueden examinar desde tres puntos de vista:

- la información que contienen
- el apoyo para adquirir y sistematizar la información
- las facilidades para recuperar información

En una base de datos terminológica, la información conceptual se codifica implícitamente a través de definiciones, contextos, indicaciones sobre dominio(s), etc. En cambio, en una base de conocimiento terminológico, esta información conceptual se codifica de forma explícita. El grado de estructura resultante que se impone en la información da lugar a tres derivaciones. Primero, permite la representación explícita de relaciones conceptuales (por oposición a las representaciones implícitas de las definiciones o contextos de una BDT). En segundo lugar, permite la consistencia: ya que los conceptos genéricos se señalan de forma explícita, por ejemplo, las definiciones de todos los conceptos coordinados deben tener el mismo término como *genus*, y ya que los subconceptos heredan las características, éstos se corresponden entre los coordinados. En tercer lugar, la representación explícita de relaciones conceptuales facilita la representación gráfica de las estructuras de conocimiento. Este aspecto está enfatizado especialmente en el proyecto Cogniterm ya que considera que las representaciones gráficas ayudan al aprendizaje, y se proporciona un tipo de mapa conceptual como recomiendan numerosos psicólogos educativos.

A diferencia de las bases de datos convencionales, una base de conocimiento terminológico como Cogniterm proporciona no solo un medio para almacenar información, sino también mecanismos para ayudar a adquirir y sistematizar la información de primera mano. Los mecanismos de herencia juegan un papel importante en este aspecto: a nivel sencillo, evitan que el terminólogo tenga que repetir la información de un nivel a otro en la jerarquía y permiten la posibilidad de hacer pruebas de tipo 'y si'. A un nivel más profundo, la herencia se puede asociar a mecanismos para indicar conflictos cuando los cambios de un nivel jerárquico afecta a la estructura de conocimiento. El mecanismo de búsqueda como el que se ha implementado para Cogniterm ayuda a la adquisición ya que permite la navegación tipo hipertexto a través de las estructuras de conocimiento necesarias para descubrir espacios de conocimiento 'compatible' para un nuevo concepto.

Otras características de la interfaz de usuario, como máscaras, la matriz de comparación de características, y un entorno gráfico desarrollado, son ejemplos de las potentes herramientas

de las que dispone un entorno de BCT, diseñadas para ayudar a los terminólogos a 'conseguir el conocimiento' a través de procesos de adquisición.

Sobre la recuperación de información, desde Cogniterm señalan que las bases de datos terminológicas convencionales están limitadas por su orientación término-concepto que se rige basándose en que al conocer un término, uno puede esperar que el sistema indique (hasta cierto punto, al menos) lo que significa, cuáles son sus sinónimos, etc.

La investigación en terminología, sin embargo, parte en muchos casos del concepto al término. Por ejemplo, la terminología de la 'vida real' se basa en preguntas como: "¿Cómo se llama la máquina cuya función es W?" o "¿cómo se llama el material que tiene las características físicas, X, Y y Z?". La incapacidad de las BDT convencionales para responder a este tipo de preguntas nos lleva a la proliferación de sinónimos y cuasi-sinónimos, uno de los grandes impedimentos de la comunicación en dominios especializados. En Cogniterm, se pretendía que los usuarios pudieran acceder a los datos a través de cualquier característica conceptual para determinar si el concepto que tenían en mente tenía una denominación determinada.

3.5.2 CAOS (Computer-Aided Ontology Structuring)

CAOS (Computer-Aided Ontology Structuring) es un proyecto orientado a la construcción de sistemas de conceptos, más concretamente, en forma de ontologías (Madsen et al., 2002). Su base teórica se fundamenta en la especificación de las distintas dimensiones de un concepto mediante lo que sus autores denominan especificaciones de características. Veamos la figura que proponen Madsen y colegas y que servirá de ejemplo gráfico para entender su propuesta de formalización:

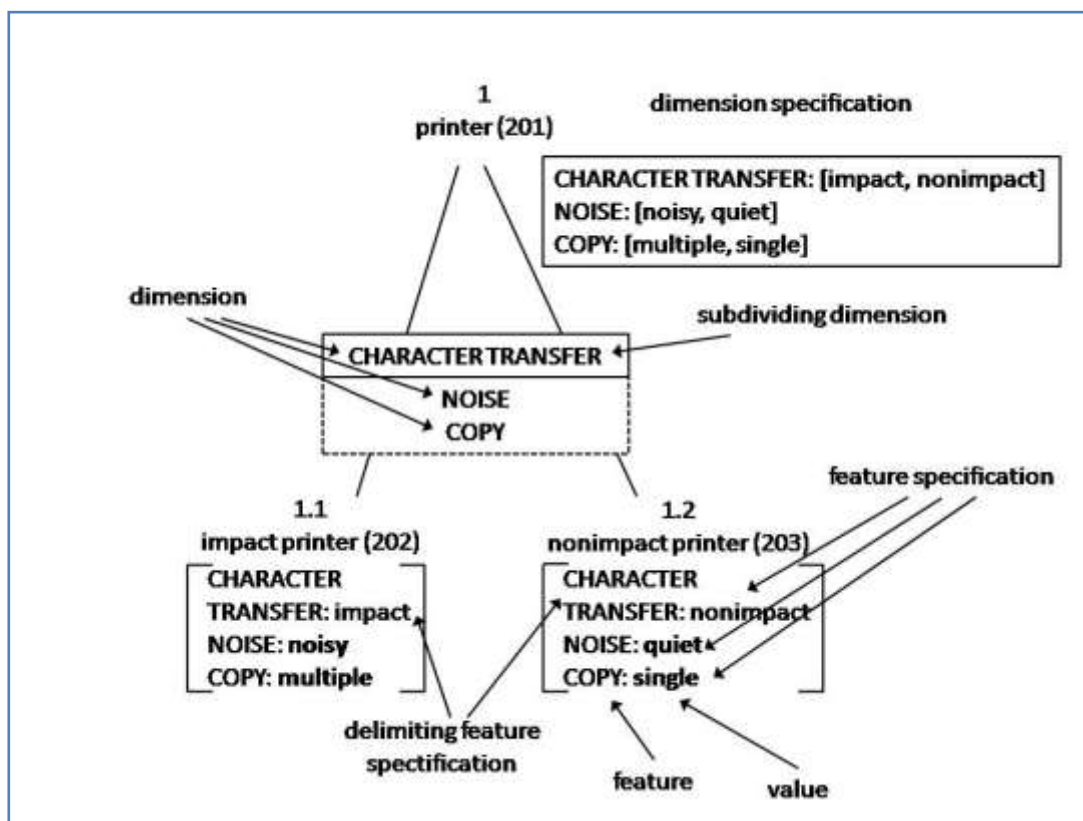


Figura 5. Sistema de conceptos en CAOS 1 (Madsen et al., 2005: 166)

En esta figura 5, los autores distinguen entre dos tipos de impresora (*impact printer*, *nonimpact printer*) que tienen tres características con valores diferentes entre ellas pero una única característica delimitadora (*character transfer*), la que se ha utilizado para establecer la subdivisión. Como señalábamos anteriormente Madsen et al. (Madsen et al., 2002) se cuestionan la necesidad de diferenciar entre características esenciales, distintivas y complementarias y proponen el binomio delimitadoras y no delimitadoras para la descripción de conceptos. Esta propuesta intenta dar una respuesta al problema de la representación de las distintas dimensiones del concepto. Dado que la especificación de características de un concepto varía según el enfoque o la vertiente con la que se mira al concepto (función, material, etc.), los conjuntos de características que lo constituyen normalmente son multidimensionales. Por tanto, un sistema conceptual será multidimensional cuando permita la representación de las distintas facetas o enfoques (Madsen et al., 2005: 170). Los autores señalan que la representación de conceptos normalmente se compone de una serie de conjuntos de pares de atributos (por ejemplo, *color*) y valores (por ejemplo, *rojo*) que Madsen y colegas denominan *tipos de especificaciones de características* y que se utilizan para delimitar y definir los conceptos; de esta forma, la representación de un sistema conceptual multidimensional debe permitir la observación de un concepto y todas las asociaciones de éste al resto según sus atributos. Sin embargo, este tipo de representación de pares de atributo y valor muchas veces resulta redundante, desde un punto de vista de delimitación de conceptos, ya que solo uno de los pares atributo-valor suele ser el causante de la distinción de conceptos. Para resolver esta redundancia y mantener la necesaria representación de las distintas dimensiones del concepto, proponen que un sistema conceptual se base únicamente en características delimitadoras siendo estas las que sirvan de criterio de subdivisión entre conceptos.

En esta propuesta, las características se codifican como especificaciones de características formales formadas por un atributo de propiedad y un valor asociado, [*Atributo: valor*]. Cuando, en la especificación de características, un atributo (no heredado) de un concepto posibilita la distinción de sus subconceptos se dice que hay una *dimensión conceptual*, además los valores asociados al atributo en la especificación de características y la dimensión del concepto constituyen una *especificación de dimensión* (*dimensión: [valor1, valor2,...]*). En el esquema de la figura 5, se puede observar la representación de una especificación de dimensión. Para diferenciar distintos conceptos en una misma dimensión conceptual se utiliza una sola característica delimitadora, en este caso el atributo *character transfer*.

3.5.3 GENOMA-KB

GENOMA-KB es un proyecto desarrollado en el IULA por el grupo IULATERM (Feliu et al., 2004). El sistema GENOMA-KB⁵ se plantea como un recurso sobre conocimiento del ámbito del genoma orientado a traductores, terminólogos y periodistas científicos (Cabré et al., 2004). Este sistema de conocimiento se compone de cuatro módulos, que se describen a continuación.

- Una base de datos textual que contiene textos especializados del dominio del genoma humano en tres lenguas: catalán, español e inglés.

⁵ Genoma-KB se puede consultar en <http://brangaene.upf.es:8080/genoma/index.jsp>

- Una base de datos factográfica y documental que contiene meta-información sobre los textos etiquetados en el corpus como información bibliográfica o metadatos relacionados con el dominio.
- Una base de datos terminológica donde se incluyen las unidades de conocimiento extraídas de los textos y se vinculan con los conceptos de la ontología.
- Una ontología donde los conceptos y unidades de conocimiento de la base de datos aparecen organizados según un conjunto de relaciones conceptuales jerárquicas y no jerárquicas.

La ontología y la base de datos terminológica se han elaborado adaptando la herramienta OntoTerm (Moreno, 2002), donde se han introducido mejoras para permitir la representación de herencia múltiple, es decir, que un concepto pueda pertenecer a más de una clase.

La ontología se construye en base a cuatro grandes categorías, que representan las categorías semánticas: eventos, objetos, relaciones y propiedades que se han ido desarrollando hasta incluir más de 5000 conceptos vinculados a través de distintos tipos de relación (equivalencia, ubicación, espació, causalidad, meronimia, etc.).

Para llevar a cabo la ontología, se ha pedido la colaboración a expertos del dominio que han proporcionado una estructura conceptual inicial de conceptos del dominio. Estos conceptos se han agrupado bajo alguna de las 21 categorías base con las que trabaja la ontología (todo, objeto, evento, propiedad, etc.). Posteriormente, el experto ha proporcionado una lista de 100 conceptos para completar la lista inicial.

En esta propuesta, la ontología funciona como recurso de ayuda para completar la base de datos terminológica. Una vez estructurados los conceptos en la ontología, éstos son descritos a través de la base de datos terminológica.

En la figura siguiente, mostramos una imagen de un extracto de la ontología para la clase ‘contagious-disease’:

The screenshot shows the Genoma KB website interface. At the top, there is a search bar with the text 'Terme de la cerca: malaltia (Català)' and 'Condició de cerca: Que contingui'. Below the search bar, the results are displayed under the heading 'Resultat de la cerca' and 'Termes trobats (25)'. The main content area is divided into three columns: 'Ontologia', 'Relacions d'hiponímia', and 'CONTAGIOUS-DISEASE'. The 'Ontologia' column shows a hierarchical structure of terms, with 'CONTAGIOUS-DISEASE' at the top, followed by a list of specific diseases: AIDS, DIPHTERIA, INFLUENZA, MEASLES, POX, and SCARLET-FEVER. The 'Relacions d'hiponímia' column is empty. The 'CONTAGIOUS-DISEASE' column is also empty. On the left side, there is a list of search results for 'malaltia', including 'malaltia contagiosa (CONTAGIOUS-DISEASE)', 'malaltia de Bornholm (EPIDEMIC-PLEURODYNIA)', 'malaltia de Christmas (HEMOPHILIA-B)', and 'malaltia de Duchenne-Griesinger (DUCHENNE-MUSCULAR-DYSTROPHY)'. The website header includes the Genoma logo and the text 'Banc de Coneixement sobre el Genoma Humà' and 'BT banc terminològic'. There are also navigation links for 'Suggeriments', 'Mapa', and 'Ajuda'.

Figura 6. Extracto de la clase 'contagious-disease' en Genoma-KB

A esta información se accede a través del buscador de su sistema que permite hacer búsquedas exactas o parciales de un término. El sistema, sin embargo, presenta inconvenientes, desde un punto de vista de su accesibilidad, ya que no nos permite o no hemos sabido obtener una lista de los términos que se han introducido en la ontología, y para acceder a la información de un concepto debemos conocer de antemano cómo se denomina. Tampoco nos permite acceder a la descripción de conceptos desde sus propiedades, a pesar de haber realizado búsquedas concretas sobre propiedades del ámbito del genoma como 'motilidad' o 'metabolismo', el sistema no nos devuelve los resultados esperados. Parece como si el sistema de búsqueda siguiera un enfoque más semasiológico que onomasiológico.

3.5.4 EcoLexicon

EcoLexicon es un proyecto desarrollado por el grupo Lexicon de la Universidad de Granada (Faber, 2009). El resultado del proyecto EcoLexicon es un recurso terminológico basado en la Teoría Basada en Marcos, donde la representación de conceptos se realiza mediante marcos conceptuales que van conformando todo un sistema conceptual de un dominio. Para ello, se basan en la derivación de un sistema conceptual de dominio mediante un enfoque integrado *top-down* y *bottom-up*. El enfoque *bottom-up* consiste en extraer información de un corpus de textos en varias lenguas, especialmente relacionado con un dominio. El enfoque *top-down* incluye la información proporcionada por los diccionarios de especialidad y otro material de referencia, que se completa con la ayuda de los expertos del dominio.

Al mismo tiempo, se va especificando el marco conceptual que subyace en un evento del dominio de conocimiento. Las categorías más genéricas o básicas de un dominio se configuran mediante una interfaz del entorno de acción o evento del dominio prototípico. Esto proporciona una plantilla aplicable a todos los niveles de estructuración de la información. La estructura que se obtiene facilita y mejora la adquisición de conocimiento ya que la información de las entradas terminológicas es coherente interna y externamente (Faber, 2009).

En esta propuesta del grupo LexiCon, la descripción de los dominios de especialidad se basa en los eventos que normalmente tienen lugar en él y se representan de acuerdo con eso.

Cada área de conocimiento tiene su propia plantilla de evento y contiene las categorías prototípicas generales que servirán de marco para organizar los conceptos más específicos. Los conceptos específicos de cada categoría se organizan en forma de red mediante relaciones jerárquicas y no jerárquicas. En EcoLexicon⁶, el sistema conceptual está conectado con los términos en distintos idiomas.

Dentro de un dominio, cada subdominio se configura a su vez mediante una plantilla que contiene las relaciones conceptuales prototípicas que se dan en cada uno de ellos. Las definiciones terminológicas se tienen muy en cuenta y se basan en los datos extraídos del análisis del corpus, que se considera la principal fuente de información paradigmática y sintagmática de los términos.

Un ejemplo que propone Faber dentro del proyecto EcoLexicon, es la definición de 'erosión'. La 'erosión' es un proceso que pertenece a la plantilla de procesos dentro del evento 'ingeniería de costas'. La información que representan es la siguiente:

- puede ocurrir en una estación concreta del año
- puede ocurrir hacia una dirección determinada
- la realiza un agente (fuerza natural)
- afecta a un punto geográfico concreto o entidad medioambiental
- produce como resultado la posible modificación de la entidad afectada

A partir del estudio de los datos del corpus (concordancias, etc.) se van identificando los atributos que ayudan a construir la definición de 'erosión' y mapear las relaciones conceptuales que se dan entre 'erosión' y otros conceptos. Asimismo, se obtiene información sobre su potencial combinatorio en una o más lenguas. Faber presenta una serie de fragmentos del corpus donde se pueden ir observando los atributos que lo describen. Por ejemplo, del fragmento de la concordancia en el corpus en lengua inglesa se obtiene información sobre que 'erosión' es un proceso (ibídem: 125):

(1) diesi is a result of soil weathering and erosion processes. Excessive loading of

Otro ejemplo de concordancias donde se observa que 'erosión' es un proceso inducido por un humano o una fuerza natural sería el siguiente:

(16) iglacial environments: nivation; <i>eolian</i> erosion and deposition; and fluvial eros
(17) ited sediments from <i>wind</i> , <i>water</i> , or <i>ice</i> erosion . Given more time, these soils wi
(18) urbances, for instance by <i>flood-induced</i> erosion , redistribution of sediment or a
(19) e most important cause of <i>human-induced</i> erosion is interruption of

⁶ EcoLexicon se puede consultar en <http://ecolexicon.ugr.es/en/>

sediment sour
(20) stal bluffs and episodic, <i>storm-induced erosion</i> of dunes and barrier beaches. A
(21) y 24 acres per year, while <i>wave-induced erosion</i> is approximately nine acres. Ac

La información básica de un concepto, como es el caso de ‘erosión’ en lengua inglesa, se activa en la creación de términos más específicos presentes en el corpus y que se consideran hipónimos del concepto descrito. Estas formas nominales complejas se entienden como proposiciones comprimidas que tienen su propia sintaxis. En su forma extendida, en el caso de los hipónimos de ‘erosión’ se observarían del siguiente modo (ibídem: 127):

Term	Verb	Agent (Arg1)	Result (Arg2)
storm-induced erosion	INDUCE	storm	erosion
human-induced erosion	INDUCE	humans	erosion
wave-induced erosion	INDUCE	wave	Erosion

Y en el sistema EcoLexicon, la consulta de ‘erosión’ proporciona el siguiente resultado (figura 7):

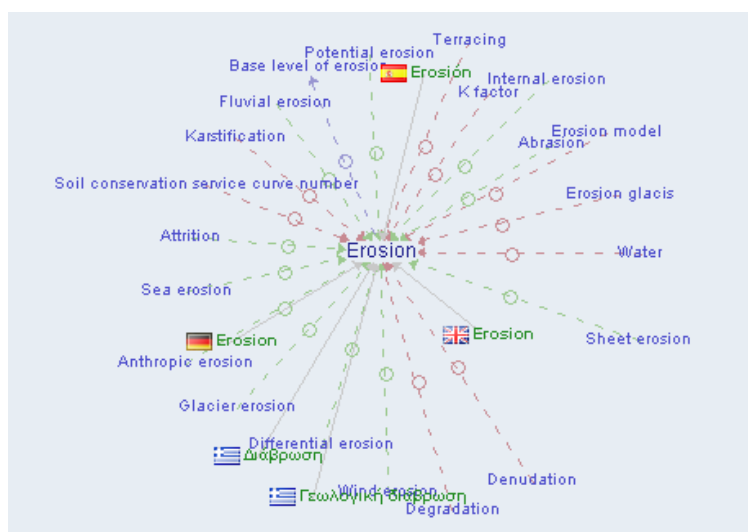


Figura 7. Resultado de la búsqueda de ‘erosión’ en EcoLexicon

Las representaciones de proposiciones se pueden activar de distintos modos, dependiendo de la lengua y las reglas para la formación de términos. Estas construcciones de argumentos proporcionan los medios básicos para elaborar expresiones de oraciones en una lengua. El ejemplo anterior, se correspondería con la construcción *X causa Y*. Cada una de estas construcciones se observa como una unidad básica de lenguaje que puede relacionarse con otras construcciones y formar una red, de forma que las estructuras semánticas queden emparejadas de la forma más general posible (*erosion y storm, erosion y humans, etc.*).

Finalmente, un aspecto que ayuda a interpretar y entender mejor el conocimiento que se plasma mediante el enfoque de la Terminología Basada en Marcos es el uso de las imágenes para representar los conceptos especializados. Por ejemplo, si consultamos la ficha de ‘precipitación’, podemos encontrar un recurso auxiliar que nos muestra una imagen relacionada con ‘precipitación’.



Figura 8. Recurso auxiliar con imágenes sobre los conceptos en EcoLexicon

3.5.5 Oncoterm

Oncoterm es otro proyecto de investigación interdisciplinar del grupo Lexicon (López et al., 2006). En este caso trata sobre terminología médica y tiene como objetivo global la elaboración de un sistema de información específico del subdominio biomédico de la oncología. Para codificar el conocimiento se basan en el Modelo Lexemático Funcional (Mignorance, 1984 y 1995) y la Semántica de Marcos (Fillmore, 1982).

En Oncoterm un concepto conecta con otros mediante hiperenlaces (López et al., 2006). El resultado es una base de datos terminológica bilingüe (español e inglés) que contiene 1896 conceptos relacionados con el cáncer y relacionados entre sí reflejando las relaciones de significado que tienen entre ellos. Esta conceptualización contiene 4033 términos en inglés y español incluyendo formas completas, sinónimos y siglas. Además, las entradas contienen otra información como definiciones, contextos, imágenes, enlaces a webs relevantes o archivos de concordancia en formato KWIC para estudios lingüísticos.

Este sistema que conjuga la ontología con la base de datos tradicional está disponible de forma gratuita en Internet⁷ y dispone de una interfaz que permite la búsqueda de términos por orden alfabético al mismo tiempo que podemos acceder a otros términos relacionados a través de enlaces (figura 9).

⁷ Oncoterm se puede consultar en <http://www.ugr.es/~oncoterm/alpha-index.html>

En GENOMA, Eco-Lexicon u Oncoterm la información se presenta a través de una interfaz de consulta que sirve para acceder a la ontología y también a otra información vinculada a la ontología. En este aspecto, hemos probado GENOMA-KB y Eco-Lexicon para probar sus búsquedas por métodos onomasiológicos y notamos que todavía pueden mejorarse. En el caso de GENOMA, como ya hemos comentado, el sistema no nos permite realizar búsquedas por propiedad o característica y en el caso de Eco-Lexicon, es cierto que al buscar una característica sí la vincula y la sitúa dentro del sistema de conceptos pero notamos ruido en las respuestas. Por ejemplo, si buscamos "quality" en Eco-Lexicon, nos devuelve el marco de "quantity" cuando por aproximación deberían aparecer antes los de "ecological quality" o "quality of life". No obstante, esto es un mero detalle técnico. Es mucho más remarcable, observar la agilidad con la que el sistema presenta los marcos y la claridad con la que están representados.

En resumen, los sistemas presentados son, en nuestra opinión, un ejemplo a seguir y del que tomar ideas en nuestra ontología.

4 Metodologías para la creación de recursos terminológicos

4.1 Metodología para la elaboración de un banco de datos terminológico

Para elaborar un banco de datos, Cabré (Cabré, 1992: 385) recomienda seguir una serie de pasos que principalmente son:

- Elaborar un proyecto que defina a grandes rasgos el banco de datos indicando, por ejemplo, las necesidades que va a cubrir el banco de datos, sus destinatarios, las prioridades en caso de conflictos, obstáculos que puedan aparecer, especialistas y usuarios que formarán el equipo de trabajo, tareas que debe realizar cada miembro del equipo, plan de trabajo y recursos necesarios.
- Realización de un estudio de viabilidad. Este estudio permitirá identificar las soluciones adecuadas en cuanto a objetivos y funciones del banco de datos, aspectos informáticos, dimensiones, organización, ubicación, difusión y rentabilidad.
- Concepción general del banco de datos. En este punto será necesario describir la estructuración general del banco de datos y del proceso funcional de los datos. Entre otras tareas, destaca la formación del equipo y el análisis de los aspectos generales del banco, como pueden ser ficheros a utilizar o compatibilidad de sistemas.
- Selección del ordenador y del software. Se pueden plantear dos opciones en este punto, utilizar un software comercial ya existente o utilizar un software a medida, creándolo especialmente para cumplir con los propósitos del proyecto. Cabré señala que existen ventajas e inconvenientes en ambas opciones, entre las desventajas destaca el coste de los productos comerciales y el tiempo que hay que invertir en el caso de los hechos a medida. En las ventajas, el software comercial suele ser compatible con otros sistemas y el software a medida, cumple con los objetivos concretos del proyecto.
- Concepción detallada del banco de datos. Este paso consiste en concretar de manera explícita las características de las entradas del banco de datos, su sistema de almacenamiento y la salida de información.
- Creación de un proyecto piloto o prototipo. Se creará el banco de datos en forma de proyecto piloto, se realizará el análisis y programación así como la redacción de manuales y protocolos, entrada de datos, verificación, etc.
- Creación del banco. Si la experiencia anterior es positiva.
- Actualización. Esta fase es permanente e inevitable si se quiere que el banco de datos siga siendo útil.

4.2 Metodologías sobre elaboración de bases de conocimiento terminológico y ontologías

A continuación, presentamos algunas metodologías utilizadas en distintos proyectos de investigación para elaborar sistemas de representación de conocimiento. Como hemos observado en la recapitulación del capítulo §2.4.6 la distinción entre base de conocimiento y

ontología es difusa y, a continuación, veremos cómo autores que hablan de un mismo sistema, intercambian estas dos denominaciones restando importancia a su distinción. No obstante, es necesario mencionar la observación de Condamines (2003: 57) que señala que la semejanza entre una base de conocimiento terminológico y una ontología se da cuando éstas últimas se elaboran considerando el lenguaje presente en corpus o su uso y que su principal diferencia subyace en el enfoque predominante que será el de la Ingeniería del Conocimiento, en el caso de las ontologías o de la terminología y la lingüística, en el caso de las base de datos de conocimiento.

4.2.1 Metodología para elaborar una base de conocimiento terminológico

De acuerdo con Meyer (Meyer et al., 1992a), la creación de bases de conocimiento terminológico (BCT) implica tres actividades principales que conllevan sus correspondientes observaciones y deben tenerse en cuenta a la hora de establecer una metodología de trabajo:

- La adquisición de conocimiento. De acuerdo con Meyer la adquisición de conocimiento de un dominio puede tener lugar bien a partir de expertos del campo o bien a partir de textos. Para ello recomienda tener en cuenta que la experiencia en un dominio se basa en tres elementos: realización, entendimiento y comunicación. Siguiendo a autores como Gaines (Gaines, 1990), Meyer enfatiza el difícil papel que debe jugar el experto en la creación de una BCT ya que se le pide que sea practicante del dominio, científico y profesor. Según Meyer, los expertos no siempre son buenos profesores, pudiendo, por ejemplo, no saber expresar bien un conocimiento. Además, su comprensión del campo puede variar, lo que derivaría en presentar al terminólogo/ingeniero del conocimiento problemas de inconsistencia y contradicciones.
- Formalización de conocimiento. La siguiente actividad consiste en organizar el conocimiento de un dominio (sus conceptos y sus relaciones) de una manera formal y estructurada. En este punto, Meyer señala que la metodología de trabajo debe tener en cuenta que el conocimiento no viene preparado para ser usado ya menudo puede ser contradictorio o inconsistente (Meyer et al., 1992a:957). Además, las representaciones frecuentemente tienen un componente multidimensional, ya que el entendimiento de un sistema conceptual dependerá del punto de vista del experto. Con ello, Meyer enfatiza la dificultad de ‘captar’ el conocimiento de un dominio debido a los cambios constantes y a que el conocimiento emergente puede ser incompleto o no estar claro. Desde el punto de vista del terminólogo/ingeniero del conocimiento existen distintos grados de ‘claridad’ y ‘profundidad’, dado que la adquisición del conocimiento incrementa, algunos conceptos son más claros o se entienden más profundamente que otros en un determinado momento.
- Refinamiento de conocimiento. Una vez formalizado, el conocimiento se debe refinar de dos modos: 1) validándose probando el sistema en la aplicación específica y/o 2)

debe actualizarse, por ejemplo, a medida que el conocimiento del terminólogo/ingeniero del conocimiento crece, cuando el dominio cambia. De nuevo, el refinamiento de conocimiento implica la adquisición y la formalización de éste, lo que convierte el ciclo de Ingeniería del Conocimiento en un proceso continuo.

A la hora de decidir una metodología preliminar para su trabajo, Meyer recurre en un primer momento a la bibliografía sobre lexicología computacional, y la Ingeniería del Conocimiento, lo que en sus palabras resulta poco exitoso. Meyer reprocha que en los grandes proyectos de adquisición de conocimiento resulte difícil encontrar una guía metodológica. Por otra parte, señala que en el campo de la terminología, las metodologías se basan esencialmente en aquello que se usa tradicionalmente en terminología (Sager, 1993), un punto de vista razonable si se tiene en cuenta que aunque los terminólogos no acostumbraban a crear BCT, el análisis conceptual ha sido una parte central de su trabajo.

Los terminólogos son conscientes de la importancia de profundizar en el conocimiento del dominio y utilizan metodologías de trabajo que se parecen a las técnicas de análisis conceptual que aparecen en la bibliografía de la Ingeniería del Conocimiento. Por ejemplo, cuando se habla de la descripción de las características de un concepto a través de pares de atributo-valor o de la construcción de redes conceptuales (Meyer et al., 1992b:169).

La metodología que proponen se inicia con una **lectura introductoria del dominio**. A través de esta lectura, el terminólogo debe identificar las **principales relaciones conceptuales** con el fin de establecer los límites del dominio e identificar posibles subdominios, los más relevantes se analizarán con más profundidad.

Se establece una **plantilla de características conceptuales** para el subdominio, que se usa como guía para la adquisición del conocimiento y se heredará en la jerarquía conceptual, donde también habrá especializaciones.

La **información lingüística** se introduce en el sistema a medida que se va recogiendo (principalmente a partir de un corpus). Un concepto se integrara en la jerarquía siempre que se conozca su superconcepto, en caso negativo o de duda, el concepto se clasifica bajo la etiqueta 'no clasificado' (éstos puede aparecer en distintos niveles de la jerarquía).

Se construyen **definiciones intensionales** con ayuda de la matriz de comparación de características y se repiten los mismos en los subdominios, hasta que se completen.

Además de esta breve metodología, Meyer et al. (Meyer et al., 1992b) analizan una serie de aspectos que resultan determinantes en el proceso de elaboración de un banco de conocimiento terminológico.

Las rutas de adquisición del conocimiento. Uno de los problemas que han encontrado al aplicar esta metodología es que la adquisición de conocimiento no es un camino recto con un objetivo final 'visible'. A pesar de seguir metodologías de la terminología tradicional con respecto a una adquisición orientada al subdominio y con un enfoque *top-down*. Meyer et al. (ibídem), señalan que, a menudo, han tenido que desviarse del subdominio principal cuando se han encontrado términos relacionados con otro subdominio vecino u otro campo, y entonces trabajar desde una perspectiva *bottom-up* en combinación con la *top-down*. Eso se une a que han

observado que mientras en los subdominios que se han tratado predominan las relaciones genéricas-específicas y las relaciones parte-todo, en la bibliografía sobre Wordnet, por ejemplo, se observa que en los subdominios vinculados a otras clases semánticas las relaciones presentes en el análisis son de otro tipo (Miller y Fellbaum, 1991).

La multidimensionalidad. Aunque los terminólogos son conscientes de que un determinado dominio puede dividirse de distintas formas, no existe una tradición para reflejarlo de una forma seria dada la dificultad que comporta hacerlo mediante técnicas basadas en sistemas tradicionales. Según Meyer, algunos de los problemas que surgen están relacionados con el efecto de la multidimensionalidad en las rutas de adquisición y en cómo la tecnología puede ayudar mejor al mantenimiento de la claridad conceptual a medida que el número de dimensiones crece (por ejemplo, a través de herramientas de filtrado como las implementadas) o cómo se puede reflejar la multidimensionalidad en la construcción de definiciones.

La validación. Meyer et al. (Meyer et al., 1992b: 103) entienden que la validación por parte de expertos u otros terminólogos es una parte importante del trabajo terminológico. Sin embargo, al mismo tiempo aceptan que en el enfoque presentado esto resulta algo complicado ya que la BCT es como un hipertexto, y requiere técnicas de revisión que van más allá de lo comúnmente aplicado por expertos. Es necesario, por tanto, investigar en qué puntos se requiere la consulta de validadores, que técnicas de elicitación se pueden utilizar en cada paso o cómo se puede tratar la inconsistencia de opiniones.

El aumento de la automatización. Los esfuerzos de investigación están orientados a facilitar (y no automatizar) el proceso de adquisición de conocimiento para desarrollar e implementar el concepto de BCT. Esto es algo inherente a la mayoría de proyectos de adquisición. De acuerdo con Meyer, a medida que el concepto de BCT se aclare, los investigadores en BCT y bancos de conocimiento lexicográfico colaborarán para explorar posibilidades para automatizar el enfoque de la adquisición

4.2.2 Metodología de la Teoría Basada en Marcos

El grupo de Faber, propulsor de la Teoría basada en Marcos (TBM), ha desarrollado distintos proyectos (López et al., 2009) de los que se desprende una metodología de trabajo acorde con su marco teórico. En este caso, vamos a profundizar en el proyecto OncoTerm.

El proyecto **OncoTerm** se centra en la terminología médica y consiste en la elaboración de un sistema de información específico del subdominio biomédico de la oncología (Faber, 1999; Faber, 2002; Faber 2010). Para ello clasifican los términos y los representan en una base de datos que permite definir atributos, hipónimos, hiperónimos y relaciones entre conceptos. En su análisis de conceptos de la oncología detectan sondas de información que servirán para definir y delimitar los conceptos a partir de textos especializados y obras de referencia. En concreto, en el caso de las características de un concepto y su representación, en este trabajo no se parte de ninguna clasificación de características previa, sino que se establece un enfoque *bottom-up* que parte del material para establecer los elementos necesarios aunque se intenta representar el máximo de características intensionales y extensionales necesarias para definir el concepto.

El proyecto OncoTerm, parte de un enfoque donde prima la gestión terminológica basada en el conocimiento. En este enfoque los términos están asociados a un marco conceptual de referencia independiente de la lengua. El marco conceptual se plasma en una ontología,

entendida como un repositorio de conceptos que incluye entidades, sus propiedades, funciones y limitaciones, los procesos a los que afectan estas entidades y relaciones entre entidades. El sistema conceptual queda reflejado en la definición terminográfica de cada unidad terminológica y la definición terminográfica está basada en la extracción de información de un corpus textual, en la consulta a diccionarios y expertos y en el análisis de imágenes.

La metodología utilizada en OncoTerm no se compone de fases lineales sino que consiste en un conjunto de actividades que se suceden a veces de forma paralela y se retroalimentan:

- Recopilación de corpus en soporte electrónico
- Lista de conceptos y términos
- Creación de un gestor de información
- Elaboración de un marco conceptual de referencia mediante la elaboración de jerarquías terminográficas
- Elaboración de entradas terminológicas

El corpus en soporte electrónico se compone de textos extraídos de Internet, de enciclopedias, manuales y publicaciones médicas en formato CD-ROM y de manuales en papel que fueron escaneados. En este proyecto trabajan con el inglés y el español así que recopilaron un corpus para cada lengua.

La extracción de una lista de conceptos y términos se realiza a partir del corpus utilizando herramientas de análisis textual. A través de listas de frecuencia y líneas de concordancia se identificaron los términos de la lista y algunas relaciones conceptuales entre estos.

El gestor de información utilizado fue OntoTerm (Moreno, 2002), desarrollado expresamente para el proyecto. En este sistema se introducían los conceptos con la ayuda de un editor de ontologías implementado en OntoTerm y por otra parte, un segundo módulo servía para editar la base de datos terminográfica, donde se elaboraban las entradas terminográficas vinculadas a los conceptos de la ontología.

Para elaborar un marco conceptual que se representara en el gestor de información se establecieron jerarquías de términos mediante el análisis de definiciones lexicográficas y del corpus. Mediante este método se fueron identificando las grandes categorías conceptuales que componían el sistema, por ejemplo, PROCEDIMIENTO, TUMOR, SÍNTOMA o MEDICAMENTO. Estas categorías se engloban dentro de lo que Faber denomina *EVENTO MÉDICO ONCOLÓGICO* (Faber, 2002: 8). La estructura interna de cada categoría conceptual queda representada por un conjunto de tipos de información sobre las características de un concepto (intensión) y sobre las entidades que pueden pertenecer a esa categoría (extensión). A partir de estas categorías se va modelando la ontología, insertando cada concepto en el marco conceptual EVENTO MÉDICO Y estableciendo relaciones entre conceptos, siendo la más habitual la relación ‘tipo de’ (López et al., 2006).

Finalmente, para la elaboración de entradas terminológicas, en el proyecto OntoTerm se siguieron los siguientes pasos:

- Elaboración de definiciones coherentes e informativas a partir de la aplicación de esquemas categoriales.
- Recogida de la información gramatical básica.
- Inserción de líneas de concordancia representativas
- Creación de bancos de recursos (imágenes, URL, etc.)

La elaboración de las definiciones se lleva a cabo de forma consensuada consultando el corpus y las entradas de distintos diccionarios. En la definición, se tendrá en cuenta que aparezca el hiperónimo inmediatamente superior para, entre otros, representar su lugar en la jerarquía.

Introducidas las definiciones, se incluye la información gramatical básica y las líneas de concordancia representativas así como recursos multimedia que puedan enriquecer las representaciones lingüísticas de los conceptos.

La aplicación informática OntoTerm se ha utilizado tanto en la fase de elaboración del marco conceptual como en la de entradas terminológicas. Primero, se ha creado una base de datos terminológica, mediante la extensión de una ontología general con el módulo de edición de ontologías de OntoTerm y se han ido anidado los conceptos específicos. A continuación, usando el módulo de edición de la base de datos, se ha llevado a cabo la tarea de describir los términos. Estos términos también han sido enlazados a los conceptos de la ontología. Para realizar la implementación del sistema conceptual se han utilizado las categorías del estándar ISO 12620, que se resumen en el siguiente esquema:

ISO 12620: GRUPOS DE CATEGORÍAS DE DATOS

Categorías de datos relacionados con el TÉRMINO

Subgrupo 1: categoría de datos que contiene un término u otra información tratada como tal (p. ej. una unidad fraseológica o un texto estándar).

Subgrupo 2: categorías de datos que contienen información sobre el término.

Subgrupo 3: categorías de datos sobre la equivalencia entre términos asignados al mismo concepto o a conceptos similares.

Categorías de datos descriptivas relacionadas con el CONCEPTO

Subgrupo 4: categorías de datos que clasifican los conceptos en áreas o sub-áreas temáticas.

Subgrupo 5: categorías de datos para la descripción del concepto, por ejemplo, diferentes tipos de definiciones, explicaciones o material contextual usado con el fin de definir el concepto, o para determinar el área temática y el concepto al que se asigna un término.

Subgrupo 6: categorías de datos que indican relaciones entre dos conceptos.

Subgrupo 7: categorías de datos que se usan para indicar la posición de los conceptos en un sistema conceptual.

Subgrupo 8: categoría de datos *nota*. Esta categoría aparece aislada, ya que puede asignarse a cualquier otra categoría y, por tanto, no puede subordinarse a ningún grupo específico.

Categorías de datos ADMINISTRATIVAS (gestión)

Subgrupo 9: categorías de datos de los lenguajes documentales y tesauros.

Subgrupo 10: categorías de datos de la información y administrativa.

Aunque estos diez subgrupos engloban más de 150 categorías de datos, en Oncoterm, no se pretende que una base de datos incluya a todas, sino solo a las que se consideran necesarias para representar la información pertinente a un determinado proyecto terminológico.

Para Oncoterm, la información más interesante son las categorías de datos relacionados con el concepto (subgrupos 4-8). Con el grupo 4 se consigue asignar un concepto a un área temática. El subgrupo 7 enlaza los conceptos formando una red o estructura conceptual, establecida en el seno de un área de especialidad determinada mediante relaciones jerárquicas. El grupo 9, por otra parte, enlaza los conceptos con los términos de un tesoro o lenguaje documentario, que suele cubrir un área de conocimiento más amplia que un sistema conceptual y que se ha construido para indexar documentos y para recuperar información en bases de datos documentales.

Las ventajas de usar el estándar englobado en la norma ISO 12620 (ISO, 1999) son la versatilidad y combinación de categorías de datos que ofrece. Además, su formalismo permite expresar información terminológica de forma multidimensional, reutilizable e intercambiable.

4.2.3 **Metodología para transformar un modelo de conceptos en un modelo de datos conceptuales**

En un artículo de Madsen y Odgaard (Madsen y Odgaard, 2010: 537-544) encontramos una propuesta que distingue entre la fase en la que describimos un sistema de conceptos, a través de características y relaciones, y su posterior transformación en un modelo de datos conceptual (clases, atributos, etc.). De acuerdo con estas autoras, la fase de descripción del modelo de conceptos se omite en muchos casos en computación y esto puede dar lugar a problemas y malentendidos en el modelado de datos. Para llevar a cabo la transformación adecuadamente, las autoras proponen un algoritmo semiautomático que convierte un modelo conceptual en un modelo de datos conceptual. Este algoritmo se lleva a cabo durante un proceso de modelado que se divide en cuatro fases y que explican utilizando como ejemplo el modelado de un sistema sobre conferencias.

En la **primera fase**, proponen que los expertos del dominio esbocen el conocimiento que quieren representar e identifiquen los conceptos y términos relevantes. Esta información se extraerá también mediante especificaciones y requisitos. Después con la ayuda de terminólogos si es necesario, los expertos organizarán los conceptos visualmente en forma de modelo conceptual donde las relaciones semánticas entre los conceptos se definen e identifican. Finalmente, esta fase termina con la delimitación de los criterios de subdivisión de las categorías expresadas y las características de los conceptos de modo que resulten descripciones precisas, simples y cortas.

En la **segunda fase**, se desarrolla un modelo de datos conceptual en UML (lenguaje de modelado unificado de datos del campo de la ingeniería orientada a objetos que a través de un conjunto de técnicas de anotación gráfica permite crear modelos visuales en forma de diagramas). Las autoras explican que los conceptos presentes en el modelo conceptual desarrollado en la fase anterior, normalmente, se corresponderán con clases, atributos o valores de atributo pero también puede darse el caso de que algunos conceptos no se correspondan directamente con ninguno de estos elementos. Además, observan que las relaciones entre los

conceptos tal y como estaban en el modelo anterior y como se representarán ahora pueden variar.

El algoritmo que proponen para transformar la información obtenida en la fase 1 en un modelo de datos en UML de forma semi-automática tiene lugar en esta fase y parte de la identificación de conceptos foco. Un concepto foco es un concepto que se caracteriza por no tener ningún concepto superordinado o que no participa de ninguna relación de asociación como concepto de llegada (conceptos considerados receptores de relaciones). Por ejemplo, 'person' en caso del sistema sobre conferencias.

A continuación, se analiza un conjunto de propiedades para determinar qué rol debe tener el concepto foco, y que importancia tendrán los criterios de subdivisión y las relaciones expresadas. Por ejemplo, si un concepto foco tiene muchos conceptos subordinados, éste se convertirá en una clase UML y cada criterio de subdivisión se convertirá en otra clase que se denominará de acuerdo con el atributo que corresponda al criterio y además se creará una relación (*attribute*) de tipo "has" (tiene) entre la clase del concepto foco y la clase del criterio de subdivisión, de modo que los conceptos subordinados se transformen en valores de la relación (*attribute values*) en la clase que representa al criterio de subdivisión.

Otra parte del algoritmo que proponen sirve cuando un concepto foco tiene una relación asociativa como concepto origen. En ese caso, el concepto foco se convierte en clase (si no se ha hecho ya) y los conceptos relacionados se convierten en clases con atributos que recibirán el nombre de esas clases. Mediante una relación de tipo "has" se vincula al concepto foco con los conceptos subordinados.

Otra regla del algoritmo es que si un subconcepto de un concepto foco tiene a su vez subconceptos, o una relación como concepto origen, entonces este concepto se tratará como si fuera un concepto foco en lo que a sus subconceptos o la relación respecta.

Mediante la aplicación sucesiva de este algoritmo se consigue, según las autoras, transformar el modelo conceptual en un modelo de datos conceptual. Aun así, queda trabajo manual por hacer, como determinar la existencia de herencias múltiples o duplicados, simplificar el modelo o implementar nuevos requisitos.

La tercera fase concierne a la transformación del resultado de la fase 2 en un modelo de datos lógico. En esta fase, las clases reciben una clave primaria que se corresponde con el nombre de la clase junto con un identificador (ID) y los atributos relevantes para la clave primaria. Las relaciones de muchos a muchos se convierten en nuevas clases intermedias y se crean relaciones de tipo "uno a muchos" entre las clases originales y las nuevas. Las clases que participan en relaciones de uno a muchos se acompañan con una clave primaria en el concepto que hace de "uno" y otra clave externa para los conceptos que actuarían como receptores. En resumen, en esta fase se codifican las clases presentes en la representación y los atributos que forman las subdivisiones, reorganizado la clasificación en caso que la clasificación se pueda escalar mediante relaciones "uno a muchos".

Finalmente, la **cuarta fase** consiste en implementar el resultado de la fase anterior en un lenguaje de programación que permita realizar consultas.

4.2.4 Metodología Termontography

Hemos hablado ya del enfoque de Temmerman y la Teoría Socio-Cognitiva de la Terminología (§1.3). En este marco, Temmerman propone una metodología que combina enfoques *top-down* y *bottom-up* para capturar y representar conocimiento a partir de expertos y de textos (Temmerman, 2003: 5). La metodología *Termontography* surge en el seno del *Centrum voor Vaktaal en Communicatie*⁸ (CVC) que dirige Temmerman.

Primero, con la colaboración con especialistas del campo de dominio que va a ser objeto del análisis ontológico, se desarrolla, *top-down*, un marco de categorías y relaciones inter-categoriales. En principio, esta categorización será un patrón para la extracción manual y semi-automática de conocimiento del corpus. Sin embargo, evolucionará gradualmente en una red de relaciones semánticas más rica y desgranada. A medida que el conocimiento se vaya recuperando del material textual se irá comparando con el marco de categorías (análisis *bottom-up*).

Los resultados del análisis se reflejarán en la base de datos terminológica, que, por ejemplo, puede utilizarse como recurso de apoyo a la ingeniería formal del conocimiento.

El flujo de trabajo (figura 10) que propone *Termontography* se divide en las siguientes fases:

- Análisis
- Recogida de información
- Búsqueda
- Refinamiento
- Verificación
- Validación

La fase de análisis y la fase de recogida de información corresponden al trabajo de preparación, las restantes corresponden a la compilación de la base de datos. En la figura 10, podemos observar el flujo de trabajo según las distintas fases y cómo se relacionan entre ellas.

⁸ Página del *Centrum voor Vaktaal en Communicatie*: <http://taalkunde.ehb.be/cvc>

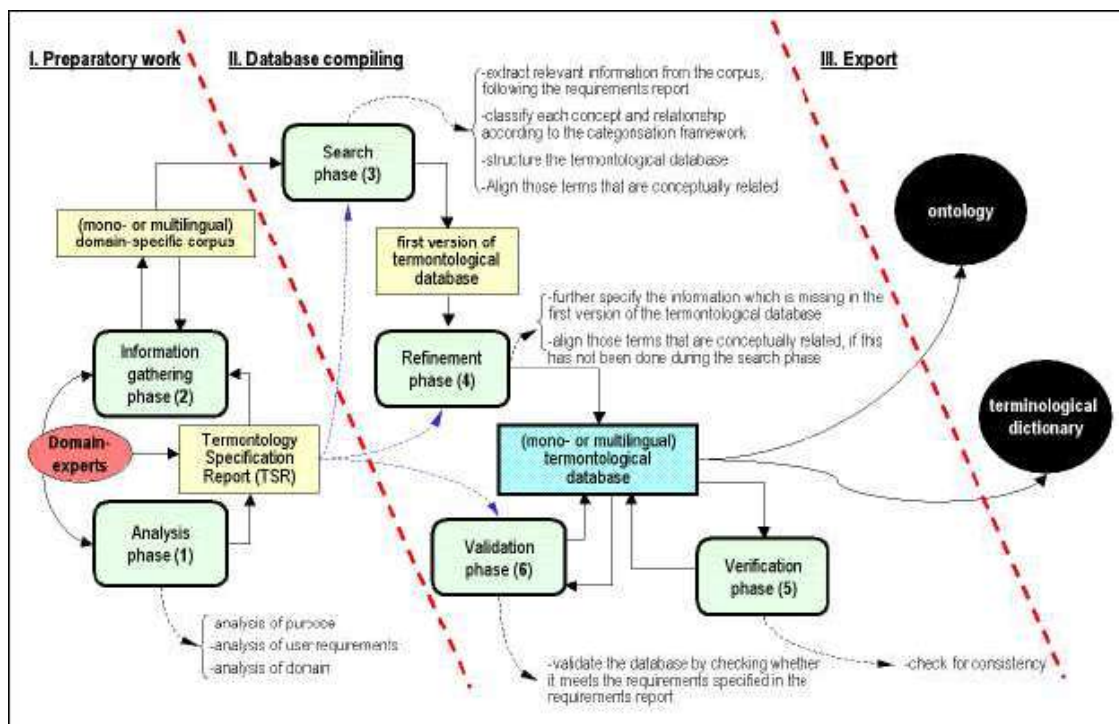


Figura 10. Flujo de trabajo en *Termonotography* (ibídem: 5)

Durante la fase de **análisis**, el terminólogo profundiza en el propósito del proyecto, el dominio y los requisitos del usuario final. Este análisis se realiza con la ayuda de expertos. En *Terminotography*, el resultado de este análisis se plasma en un informe de especificaciones terminológicas (*Terminological Specification Report, TSR*) que informa al terminógrafo sobre el contenido y el formato que debe tener la base de datos terminológica a compilar. También se elabora el marco de categorías que se utilizará durante las siguientes fases.

El marco de categorías y el TSR serán la base para realizar la búsqueda de materiales relevantes para construir un corpus del dominio. Esta fase de **recogida de información** tiene lugar también con la ayuda de los expertos.

La fase de **búsqueda** consiste en extraer del corpus términos y categorías y anidarlas en su respectiva categoría dentro del marco. En *Termonotography*, los patrones verbales que señalan relaciones entre categorías también se mapean en el marco. Según el nivel de detalle requerido se irá especificando el marco con más detalle. Del mapeado de términos y patrones verbales resultará la primera versión de la base de datos terminológica.

Para que la base de datos resulte completa, se inicia la fase de **refinamiento** donde, por ejemplo, se alinean términos equivalentes o se especifica la información faltante sobre los términos.

La fase de verificación sirve para comprobar la consistencia y adecuación de la base de datos terminológica. Esta fase viene seguida de la de **validación** que consiste en comprobar si no se ha hecho ya que todo lo realizado se adecua a lo requerido en el informe de especificación terminológica.

El trabajo se completa al exportar el conocimiento de la base de datos mono o bilingüe a un diccionario terminológico o una ontología. Distintos trabajos han utilizado esta metodología (Temmerman, 2003; Geentjens et al., 2006; Kerremans et al., 2007; Kerremans et al., 2008).

En cuanto a herramientas que permitan utilizar esta metodología destaca Termontography Workbench, que explicamos más adelante.

4.2.5 Metodología de la Ontoterminología

La metodología que se propone desde la teoría de la Ontoterminología se basa en un enfoque onomasiológico y ha sido desarrollada por el grupo Condillac (Roche et al., 2009).

De acuerdo con Roche, las metodologías que toman un **enfoque semasiológico** se acercan a la descripción del concepto desde el término, y para ello, se basan en los textos especializados. De este modo, crean representaciones basadas en lo que aparece en los textos, dando lugar a representaciones no acordes con el concepto ni con el conocimiento del dominio. Un **enfoque onomasiológico**, en cambio, no parte de los textos para elaborar una ontología sino que hace uso de expertos de un dominio para representar mediante esquemas una conceptualización de dicho dominio. Estos expertos transfieren el conocimiento del dominio de una forma independiente al lenguaje y establecen relaciones entre los conceptos no porque así se ha observado en un texto sino porque así se refleja en el dominio.

Esta metodología parte de la definición del dominio de la ontología para a continuación identificar los términos más adecuados que denotan esos conceptos (y si resulta necesario, se proponen nuevos términos normalizados para darles nombre). Roche considera que el enfoque semasiológico y el enfoque onomasiológico tienen distintos objetivos: el primero se centra en el vocabulario especializado y el segundo en la conceptualización. El resultado de ambos enfoques no se considera comparable, ya que Roche defiende que la estructura léxica que se extrae de un texto no se corresponde con la estructura conceptual (Roche, 2008:12).

Para construir una *ontoterminología* se requiere, por tanto, una metodología que va del concepto al término. Los expertos juegan un rol clave en cada paso del método. En una primera etapa del proceso, son los expertos quienes identifican los conceptos del dominio y sus relaciones. El resultado es una red conceptual semi-formal basada, principalmente, en relaciones de tipo “*es-un*” y “*parte-de*”. La jerarquía de conceptos se basa en la diferencia única, donde un concepto se define indicando el aspecto que lo diferencia de sus hermanos de clase y de su superclase.

A continuación, la red conceptual se define utilizando una herramienta propia denominada *SNCW* (*Semantic Network Craft Workbench*). El paso de convertir la representación semi-formal generada con el *SNCW* en ontología se realiza con otra herramienta *ad hoc* que denominan *OCW* (*Ontology Craft Workbench*). En este paso, la ontología se define basándose en el criterio de diferenciación específica (Roche, 2001: 25-28), que hemos explicado anteriormente en §1.5.

El siguiente paso consiste en identificar *unidades lingüísticas especializadas*, que pueden extraerse automáticamente de los textos, y redactar definiciones de éstas en lenguaje natural. Finalmente, en esta metodología se propone asociar los términos a los conceptos que se han estructurado en la ontología y agregar cualquier otra información adicional.

4.3 **Recapitulación y observaciones**

En este apartado hemos visto distintas metodologías para la elaboración de ontologías terminológicas de manera que podamos representar conocimiento terminológico. En este apartado hemos recogido un conjunto de cuestiones a las que cada metodología intenta dar respuesta con sus propuestas. A continuación presentamos estas cuestiones y comparamos las opciones que nos ofrece cada una de estas metodologías teniendo en cuenta los intereses de nuestro trabajo.

- Enfoque de la descripción y origen de los datos

En esta aspecto existen dos grandes opciones: partir de los textos del dominio como fuente para extraer terminología e información conceptual (en palabras de Roche, optar por un acercamiento semasiológico) o partir de los expertos para que elaboren un esquema o árbol conceptual que sirva como punto de partida (acercamiento onomasiológico).

Creemos que muchos autores coincidirían con Meyer en afirmar que acceder al conocimiento del dominio a través de expertos es una tarea ardua y compleja y que no siempre va a dar los resultados esperados. Como señala la autora, los expertos no tienen porqué ser grandes profesores y frecuentemente son expertos en subdominios específicos lo que puede dar lugar a discrepancias cuando tratamos de describir las relaciones entre conceptos de distintos subdominios. Además, el acceso a expertos es un asunto delicado, ya que, en nuestra opinión, muchas veces este tipo de proyectos carecen de grandes recursos económicos lo que conlleva solicitar a los expertos su tiempo de modo altruista, algo de lo que no suelen disponer la mayoría de profesionales. Enfoques como el de Faber facilitan el trabajo de extracción de información terminológica al acceder al corpus como una fuente de representación de conocimiento.

- Roles de las personas que crean la ontología o base de conocimiento

En todas las metodologías observamos la mención a expertos, terminólogos e ingenieros informáticos. En tres de las cinco metodologías que presentamos el rol de los expertos resulta clave para elaborar el mapa conceptual que se toma como punto de partida del trabajo. En el caso de los ingenieros, éstos resultan de gran importancia en la fase de formalización en un lenguaje informático y los terminólogos participan de todo el proceso de principio a fin.

Salvando las dificultades mencionadas anteriormente con respecto al uso de expertos, coincidimos en la necesidad que sean ellos quienes elaboren un primer mapa de trabajo, como señala Temmerman, a modo de categorías generales y relaciones que luego ya el terminólogo se encargará de desarrollar y detallar.

- Información que se representa en la ontología

Un punto en común en todas las metodologías es el tipo de información que se representa. Se desprende de todas estas metodologías la necesidad de que una ontología terminológica combine información conceptual con información lingüística. En este sentido, observamos que el nivel de detalle de las distintas propuestas no es comparable. Por ejemplo, en el trabajo de Faber el nivel de detalle es mucho mayor que la propuesta de Madsen, la metodología de la TBM propone describir hasta cinco tipos de información o categorías en su sistema y, en la propuesta de Madsen, prácticamente solo se elabora un mapa conceptual con relaciones.

Esto probablemente esté relacionado con el tipo de información/enfoque que prima en la metodología. En el caso de la metodología de la Ontoterminología, el esfuerzo se centra en la información conceptual y su estructuración y descripción se explican con gran nivel de detalle; en cambio, para la información lingüística la metodología se limita a proponer la inclusión de términos de uso y una definición en lenguaje natural, sin incluir contextos descriptivos, información gramatical u otro tipo de información léxica.

- Forma utilizada para representar los datos

En este apartado, hemos querido recoger metodologías enfocadas en mayor o menor grado a la representación de conceptos y términos. En la mayoría de propuestas, la representación de conceptos se elabora mediante esquemas complejos, donde se describen las relaciones de los conceptos que participan en el sistema dando lugar a ontologías. Algunas metodologías combinan el uso de ontologías para estructurar los conceptos con otros sistemas para recoger la información lingüística. En el caso de la metodología de la TBM, esto se realiza en forma de entradas terminológicas y en el caso de la metodología de la Ontoterminología, anidando esta información en un sub-apartado de la ontología.

Consideramos todas las propuestas interesantes en la medida en que consiguen vincular los conceptos a los términos. En nuestra opinión, resulta más práctico separar la información conceptual de la información lingüística a la hora de elaborar la ontología terminológica y permitir, como señala, por ejemplo, Faber, que la estructura conceptual sea independiente de la lengua en la que se esté trabajando. Lo que permite su reusabilidad en otras lenguas.

- Relaciones que se utilizan para estructurar los conceptos

La estructuración de conceptos se basa en las relaciones que existen entre los conceptos representados. Hemos observado que existen divergencias entre las distintas metodologías acerca de los tipos de relación que se deben considerar. Las relaciones más compartidas son las de tipo lógico, aunque algunos enfoques como el que se propone en el seno de la TBM abogan por representar también relaciones de tipo ontológico (principalmente de tipo “parte-todo”).

En el caso de Termontography, siguiendo el marco teórico de la Teoría Socio-Cognitiva, las relaciones que representaremos varían según el dominio. Coincidimos en este aspecto y creemos será a la luz de los resultados de la fase de análisis del dominio y considerando los objetivos del trabajo cuando podremos establecer qué relaciones vamos a representar en nuestra ontología y en qué relaciones se basará la estructura de conceptos.

- Fases de trabajo

Las fases de trabajo de las distintas metodologías resultan bastante similares a grandes rasgos. En el apartado 1.6.1 hemos querido incluir la metodología que propone Cabré para elaborar trabajos terminológicos con el fin de evidenciar las similitudes existentes entre esta metodología y las metodologías más específicas sobre ontologías.

Podemos afirmar que el trabajo se inicia con un análisis del dominio a tratar, independientemente de que este análisis se realice a partir de textos o mediante expertos del campo.

A continuación, se inicia una fase de descripción de los conceptos y sus relaciones que según la metodología se realiza mediante un acercamiento *top-down*, *bottom-up* o combinando ambos. En esta fase se utilizan textos o el conocimiento sobre el dominio para extraer la

información y se va creando una estructura de conceptos y relaciones, que en algunas metodologías ya empieza a incluir información lingüística.

La construcción de la ontología puede precederse de una fase de refinamiento (*Terminology*) o de definición semi-formal (Ontoterminología) o marco conceptual (TBM) donde se elabora una estructuración semi-formal de la ontología antes de traspasarla a un sistema informático. Esto también ocurre en el caso de la propuesta del grupo de Madsen, aunque no queda claro qué diferencia hay entre la fase de subordinación de conceptos por criterios de subdivisión (basados en una única característica delimitadora) y su posterior reagrupación en más subgrupos. Consideramos que en esta fase es donde más diferencias existen a la hora de proceder y que éstas se deben a la base teórica en la que se sustentan las distintas metodologías.

En la fase de traspaso o creación de la ontología en un sistema informático, las metodologías presentadas difieren en el sistema informático utilizado y en los pasos intermedios a realizar. En las metodologías analizadas, las herramientas que se mencionan son herramientas *ad hoc* que sirven a los propósitos concretos del trabajo a realizar.

La información lingüística suele incluirse en la fase final, una vez la conceptualización se ha llevado a cabo. Como hemos señalado anteriormente, esta información varía dependiendo de la metodología pero suele incluir los términos del campo y una definición de éstos.

En nuestra opinión, la elección de una metodología u otra dependerá del objetivo de nuestro trabajo y del marco teórico en el que decidamos sustentar el trabajo. Al mismo tiempo, creemos que algunas de las recomendaciones de estas metodologías no son incompatibles entre sí y que el trabajo de creación de una ontología de características no tiene porqué ampararse estrictamente bajo una única propuesta.

5 Herramientas para la representación de conceptos en terminología

5.1 CODE4

CODE4, denominado así por las siglas de *Conceptually Oriented Description Environment* (Skuce y Lethbridge, 1995), es una herramienta que se utilizó en el proyecto Cogniterm (Meyer, 1993). La herramienta está diseñada para ayudar a adquirir, formalizar y refinar conocimiento específico. Se plantea como una herramienta flexible, que permite a usuarios con diferentes niveles hacer uso de ella. Por ejemplo, mediante el uso de distintos registros de formalidad para acceder a la información. Además, tiene como objetivo permitir la comunicación entre distintos tipos de usuario y servir como herramienta educativa en el entorno de la representación de terminología.

Aunque se trata de una herramienta genérica, CODE4 está enfocada hacia la lingüística y especialmente en el apoyo terminológico, algo crucial para las aplicaciones de la Ingeniería del Conocimiento. Anteriormente, habían utilizado otras versiones de CODE4 orientadas a la construcción de bancos terminológicos, la ingeniería del software y el diseño de bases de datos.

Aunque no hemos podido probar la herramienta, de su bibliografía destaca su interfaz gráfica, ya que permite centrarse en partes del dominio representado (en forma de grafo) y comparar distintas sub-partes o representar distintas dimensiones así como relaciones jerárquicas y no jerárquicas. Otro aspecto a destacar son los mecanismos de control de la herencia y la consistencia que implementa la herramienta, lo que permite hacer un seguimiento de la descripción de un concepto y modificarlo en caso necesario.

El sistema de representación de conocimiento que utiliza CODE4 toma aspectos de distintos y muy variados sistemas de representación, que van desde la representación de la herencia basada en marcos, a grafos conceptuales, lógica de descripciones o representación orientada a objetos.

La unidad básica del conocimiento para CODE4 es el concepto, entendido como algo distinto al objeto que representa. Un aspecto relevante, desde el punto de vista de las características del concepto, es que en este sistema los rasgos que describen a los conceptos se denominan *properties* y se tratan como un tipo de concepto, a diferencia de otros sistemas de representación basados en marcos que no lo hacen.

Asimismo, normalmente los sistemas de representación tienen en cuenta la asociación entre una *property* (entendida como característica) y el *slot* particular que toma en un marco y lo consideran en forma de lo que viene a denominarse *triple*. Sin embargo, en este caso esto se realiza a través de *statements* (enunciados) compuestos por, como mínimo, un sujeto (un concepto) y un predicado (que codifica la *propiedad* del concepto y le atribuye un valor en caso de que sea necesario). Los enunciados en CODE4 también se tratan como si fueran conceptos que enlazan *propiedades* a *conceptos*. Los enunciados se tienen en cuenta en las jerarquías de herencia, así como en la herencia de propiedades, además, un enunciado puede tener enunciados sobre él mismo.

En total, CODE4, además de los enunciados y las propiedades, tiene otros dos tipos de conceptos primitivos, que es como ellos denominan a los elementos principales del sistema: los *meta-conceptos* y los *términos*

Los *meta-conceptos* hacen referencia información meta-lingüística sobre otros conceptos, por ejemplo, el nombre del autor que ha creado un concepto en el sistema.

Los *términos* son cada una de las denominaciones de un concepto, pudiendo ser sinónimos o no. A la hora de crear los conceptos, el sistema distingue entre conceptos descritos mediante propiedades primitivas y conceptos definidos por el usuario que se considerarán subordinados de los anteriores o de otros conceptos creados por el usuario. Los conceptos pueden ser tipos o instancias. Según el ejemplo del autor, una propiedad primitiva sería ‘fabricante de’. Y ésta aplicada por un usuario podría derivar en el siguiente enunciado: ‘fabricante de coches’ es ‘fabricante de’ ‘coche’. Este enunciado sería de tipo usuario y daría lugar a un concepto tipo. Una instancia de ‘coche’ podría ser ‘mi coche’ que, a su vez, se podría utilizar en el enunciado ‘Volkswagen’ ‘fabricante de’ ‘mi coche’.

Las descripciones de conceptos se pueden representar mediante grafos como enlaces entre el concepto que hace de sujeto y el que hace de valor, siendo el enlace el nombre de la propiedad que los vincula.

En lo que se refiere a la disposición de los conceptos en el sistema, podemos encontrarlos en forma de estructura jerárquica, de acuerdo con la relación ordenado –subordinado. Resulta interesante mencionar que este sistema permite tener múltiples superordinados.

Además, el sistema tiene en cuenta una jerarquía de propiedades o predicados. Donde las propiedades pertenecen al superconcepto ‘propiedad’ y se organizan de forma jerárquica mediante categorías a discreción del usuario. En este caso, una propiedad también puede tener más de una superpropiedad o categoría. Las dimensiones conceptuales que se generan al estructurar los conceptos basándose en alguna de estas propiedades se visualizan en CODE4.

Otra forma de visualizar los conceptos introducidos en el sistema es a través de las jerarquías de los enunciados. Una base de conocimiento tiene una única jerarquía de propiedades pero tiene muchas jerarquías de enunciados (una por cada concepto). Esto se debe a que cada concepto de la base de conocimiento hereda una sub-jerarquía de la jerarquía de propiedades, la que corresponda al concepto. Con ello, se hereda la jerarquía de enunciados que deriva de esta propiedad con respecto a sus superpropiedades, es decir, si decimos de coche:

Propiedades>entidades relacionadas>partes>partes eléctricas>partes de ignición

Estaríamos diciendo que ‘coche’ tiene ‘partes de ignición’ pero también que ‘coche’ tiene partes eléctricas.

La jerarquía ISA y la jerarquía de enunciados se utilizan para hacer inferencias sobre otros enunciados. Esto permite que en el sistema podamos buscar información a través de relaciones de tipo ISA pero también a través de enunciados.

Otras jerarquías que permite CODE4 son la jerarquía de relaciones y la jerarquía de facetas (*facets*).

La jerarquía de relaciones se construye a través de cadenas de conceptos con valor de enunciados que son una o más propiedades.

Finalmente, la jerarquía de facetas hace referencia a enunciados secundarios que representan información sobre un elemento de un enunciado y que también se pueden estructurar en forma de jerarquías. Por ejemplo, sobre el enunciado «mi coche tiene 15 años» podemos tener la faceta *valor (15 años)* pero además la faceta ‘estatus’ u otra meta-información sobre el enunciado. Los *facets* más utilizados en CODE4 son *comentario, estatus, referencia de conocimiento, modalidad y cuantificación*.

En lo que se refiere a la gestión de la herencia, el sistema permite el uso de operadores lógicos de tipo ‘and’ (lógica de descripciones).

Desafortunadamente, esta herramienta ya no está al alcance de la investigación debido a que el sistema informático para la que se creó se ha quedado obsoleto y no funciona en sistemas operativos más modernos.

5.2 Ontoterm

Ontoterm es un sistema de gestión terminológica desarrollado en la Universidad de Málaga por Antonio Moreno Ortiz (Moreno, 2002). Este sistema se ha utilizado en el proyecto Oncoterm (§3.5.5) y en el proyecto Genoma-KB (§3.5.3).

Ontoterm consta de dos módulos fundamentales. El primero, denominado *Ontology Editor* (fig. 11), es donde tiene lugar la representación conceptual propiamente. La creación de esta aplicación se inspira en el software y ontología desarrollados en el Computing Research Laboratory (CRL) de la New Mexico State University para el proyecto de traducción automática basada en el conocimiento Mikrokosmos (Beale et al., 1995). El segundo módulo (fig. 12), llamado *TermBase Editor*, está enfocado a la gestión de información terminológica. Aquí el dominio conceptual modelado en el anterior adquiere su dimensión lingüística.

La función del editor de ontologías es capaz de gestionar bases de conocimientos jerarquizadas con relaciones complejas y se integra con otros módulos de gestión de información léxica y terminológica a los que suministra información. Permite crear bases de datos nuevas y también reutilizar recursos anteriores. Un ejemplo de integración de un recurso en una nueva base de datos puede verse en Pérez Hernández & Moreno Ortiz (2000).

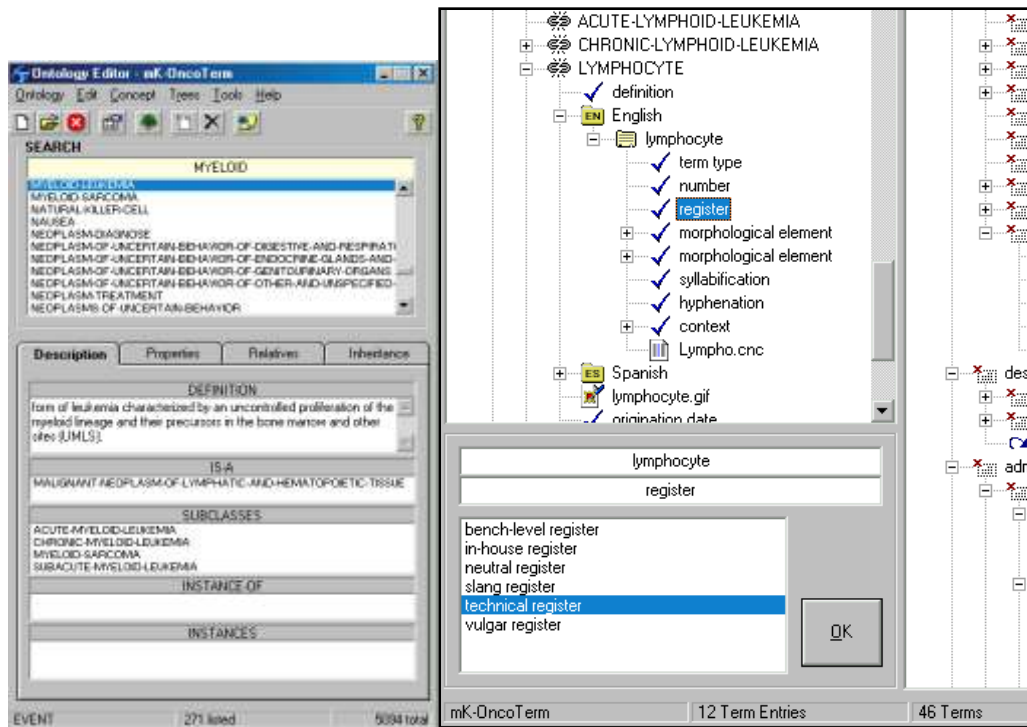


Figura 11. Módulo de la ontología Figura 12. Ejemplo de pantalla de edición en el módulo de la base de datos

La interfaz de OntoTerm (fig.13) ha sido diseñada para facilitar al usuario no experto la interactividad con las estructuras de datos de una forma lo más gráfica posible. Para conseguir esto, se utilizan controles estándar de Microsoft Windows y se hace uso de los menús de contexto. Todo, excepto la entrada de caracteres, se puede llevar a cabo empleando únicamente el ratón. También existen teclas de métodos abreviados para las tareas más comunes. No obstante, la interfaz no sirve únicamente como medio para facilitar las tareas de edición comunes, sino que también se encarga de aplicar un buen número de restricciones sobre los datos.

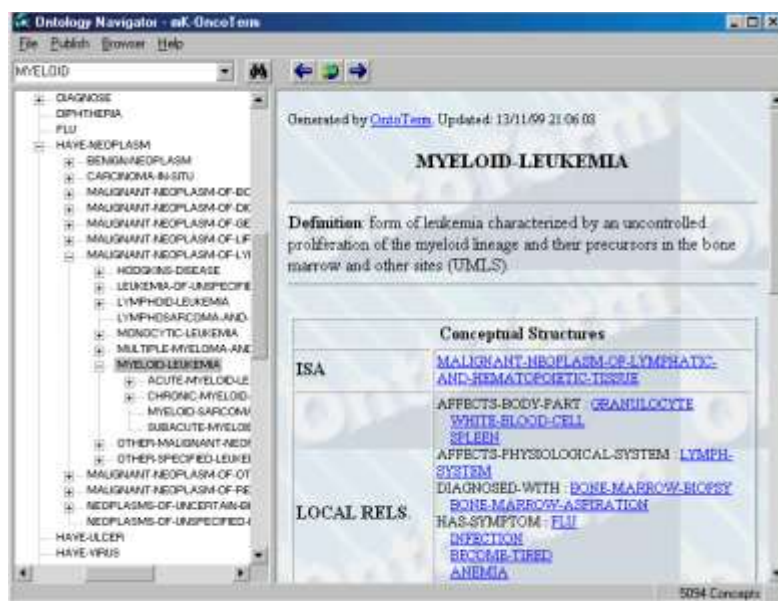


Figura 13. Muestra de la interfaz de Ontoterm

Así mismo, OntoTerm incluye una serie de herramientas destinadas a facilitar tareas de uso común, tales como la visualización gráfica de las jerarquías conceptuales, la gestión de usuarios múltiples mediante la implementación de subconjuntos de conceptos, o la revisión y publicación de información mediante un generador de informes, que permite publicar total o parcialmente la ontología en formato HTML, convirtiendo las relaciones entre conceptos en hiperenlaces.

5.3 Termontography Workbench

Termontography Workbench es un software desarrollado por el Centrum voor Vaktaal en Communicatie (CVC) que dirige Rita Temmerman en la universidad *Erasmushogeschool Brussel*. Este software nace para cubrir las necesidades de recogida, gestión e investigación de terminología sobre legislación fiscal de un proyecto multilingüe desarrollado por el CVC y se ha aplicado posteriormente en otros proyectos del grupo (de Baer et al., 2007: 2).

Este software permite gestionar un corpus (multilingüe) y extraer terminología manualmente del corpus. Además, permite crear un modelo conceptual y enlazar terminología (multilingüe) a los conceptos. Para facilitar el desarrollo del software se utilizó la API *Protégé*⁹, una herramienta de software libre que permite almacenar y gestionar modelos de datos conceptuales. Mediante *Termontography Workbench* (figura 14), un mismo terminólogo puede implementar la ontología desarrollada por expertos en el campo, construir un corpus multilingüe y anotar la terminología relevante.

⁹ <http://protege.stanford.edu/doc/pdk/kb-api.html>

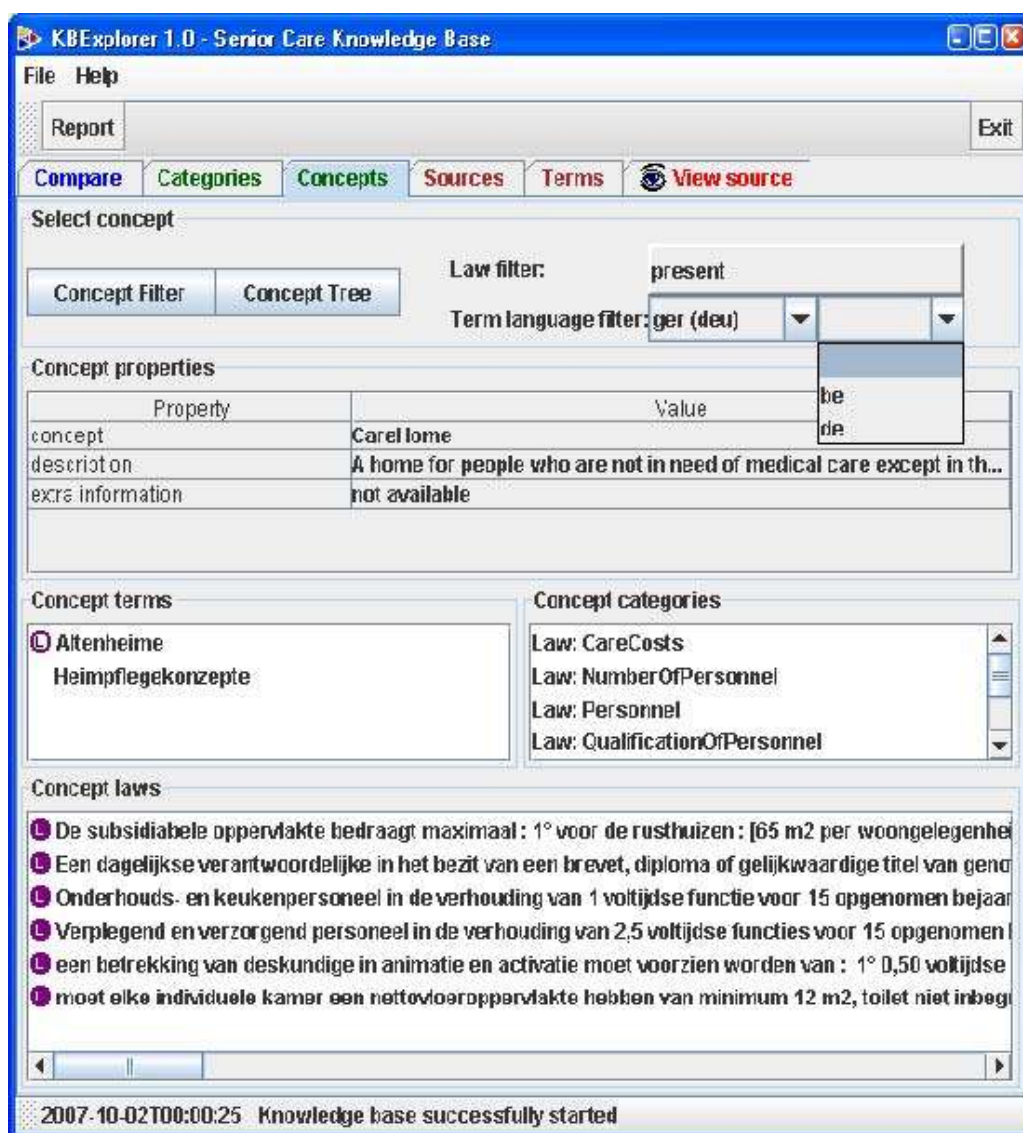


Figura 14. Interfaz de *Termonotography Workbench*

En esta herramienta encontramos los siguientes elementos (de Baer et al., 2006):

1. Fuente. Un corpus en forma de documento de texto
2. Término. En forma de cadena de texto y en una lengua.
3. Descripción (del término): descripción de un término específico
4. Relación entre términos: relación entre dos términos, por ejemplo, la forma completa de una abreviación
5. Concepto: unidad relevante de conocimiento dentro del dominio de interés
6. Relación conceptual: una relación entre dos conceptos, por ejemplo, una competencia necesaria para cierta función.
7. Categoría: una unidad de conocimiento que se utiliza para clasificar las fuentes, términos, descripciones de términos, relaciones entre términos, conceptos, relaciones conceptuales y otras categorías.

Termonotography Workbench se complementa con otra herramienta del CVC denominada *Multilingual Categorization Framework*, que explicamos a continuación y con un editor que permite gestionar los recursos terminológicos que se crean con *Terminotography Workbench* o con el *Categorization Framework Editor*.

5.3.1 Multilingual Categorization Framework Editor

Esta herramienta también se ha creado dentro del grupo *Centrum voor Vaktaal en Communicatie* (CVC) que dirige Temmerman en la Universidad de Erasmushogeschool Brussel. *Multilingual Categorization Framework Editor* es una herramienta informática para estructurar el conocimiento de un dominio que permite a los usuarios crear ontologías lingüísticas o marcos de categorías.

Esta herramienta surge de la necesidad del grupo de estructurar recursos terminológicos y lexicológicos multilingües y pretende suplir los problemas que se generaban a partir de otra herramienta del grupo denominada *KB Explorer*. *KB Explorer* se diseñó para realizar proyectos de ontologías con terminología multilingüe y hacía uso de Protégé-API para gestionar la información. El problema principal que pretende paliar *Categorization Framework Editor* es facilitar la asignación de terminología multilingüe a un mismo concepto y su recuperación y almacenamiento. Además, se han introducido mejoras respecto a la cantidad y tipología de información que podía incluirse sobre un concepto, ya que según los autores, al hacer uso del API de Protégé las posibilidades eran más limitadas y complejas (de Baer et al., 2007: 2).

La principal característica de la metodología propuesta para representar ontologías con esta herramienta es que propone construir ontologías haciendo uso únicamente de términos (multilingües), es decir, sin otra expresión lingüística que no sea un término. El *Categorization Framework Editor* se basa en 8 elementos para describir un modelo conceptual:

1. Cadena de lenguaje

Es una cadena de caracteres que se clasifica mediante una o más categorías. Cada cadena de lenguaje está clasificada mediante una categoría de lenguaje que sirve para indicar en qué idioma está la cadena de lenguaje.

2. Categoría

Las categorías se utilizan para clasificar elementos del CF. Por ejemplo, cadenas de lenguaje. Cuando una categoría hace referencia a un contexto, se denomina *meta-categoría*.

3. Meta-categoría

Las meta-categorías sirven para especificar el contexto en el que se agrupan otras categorías. Por ejemplo, cuando se agrupa la categoría '*fr*' bajo la meta-categoría '*language*', se está indicando que esa categoría representa al idioma francés.

Mediante estos tres elementos del CF: meta-categoría, categoría y cadena de lenguaje, se contextualiza la terminología. Para resolver problemas de polisemia en una misma lengua, se utilizan también meta-categorías y categorías. Pero además se utiliza otro elemento denominado *término*.

4. Término

El elemento término tiene una referencia a una cadena de lenguaje y a un significado expresado mediante una meta-categoría y una categoría de modo que distintos términos pueden

apuntar a un mismo significado. Por ejemplo, el término tiene una referencia a una cadena de lenguaje ‘*bow*’ que tiene una categoría ‘*bow*’ agrupada bajo la meta-categoría ‘*weapon*’.

De este modo ‘*bow*’, entendido como arma, se distingue de otros significados de ‘*bow*’ como el que se agrupa bajo la meta-categoría ‘*verb*’ o bajo la meta-categoría ‘*part of a ship*’.

5. Significado

El elemento significado está subyacente en las meta-categorías y/o en las categorías del CF. El significado está vinculado con una lista de términos y apunta a una meta-categoría y a una categoría. Por ejemplo, el término ‘*verb*’ hace referencia a un significado que podría reflejar mediante una meta-categoría ‘*verb*’ y en una categoría ‘*verb*’. Al enlazar un término a una (meta)categoría mediante el elemento subyacente significado, el significado de la (meta)categoría se explicita.

De este modo se pueden representar conceptos mediante terminología. Podemos identificar relaciones entre conceptos utilizando también las meta-categorías. Por ejemplo, la categoría ‘*has meronym*’ puede utilizarse para representar relaciones parte-todo. Para crear relaciones entre significados y/o términos, el CF incluye relaciones los elementos *relación bidireccional* e *instancia de relación bidireccional*.

6. Relación bidireccional

Este elemento se utiliza para indicar una relación y también puede indicar una relación inversa. Para ello, se crearía la meta-categoría ‘*relation*’ y siguiendo con el ejemplo, las categorías ‘*has meronym*’ y ‘*has holonym*’. Una relación bidireccional podría hacer referencia a ambas relaciones opuestas [*has meronym*; *has holonym*].

7. Instancia de relación bidireccional

Este elemento refleja una relación bidireccional entre dos elementos de significado o dos elementos términos. Como la dirección de una relación bidireccional suele ser relevante, la instancia de relación bidireccional distingue entre la fuente y el destino de la relación. Por ejemplo, una instancia de relación bidireccional hace referencia a la relación bi-dirección [*has meronym*, *has holonym*], la categoría fuente ‘*arm*’ y la categoría destino ‘*lower arm*’, de modo que se indica que ‘*lower arm*’ es parte de ‘*arm*’.

Mediante las instancias de relación bidireccional se puede estructurar el CF de una manera flexible y genérica. Mediante una lista de meta-categorías se puede crear una taxonomía que servirá para navegar a través del CF. Una meta-categoría puede tener varios padres de modo que hay numerosos puntos para acceder a las categorías subyacentes. Además, al proporcionar un contexto a las categorías, la jerarquía de meta-categorías tiene más funciones que un mapa de temas convencional.

La relación jerárquica género-especie se puede implementar utilizando una relación bidireccional entre categorías y se pueden añadir todos los tipos de relaciones conceptuales utilizando estas relaciones bidireccionales para construir un modelo conceptual.

Finalmente, en el CF encontramos un octavo elemento que tiene la función de añadir información adicional sobre los elementos del CF y se denomina *propiedad*.

8. Propiedad

La propiedad hace referencia a un atributo y un valor de un determinado tipo. Las propiedades se añaden sobre cada elemento del CF. Un atributo se implementa como una categoría agrupada bajo la meta-categoría '*attribute*'. Cada atributo hace referencia a un tipo de valor. La lista de tipos de valor posibles dependerá de las necesidades de implementación del CF, aunque como mínimo tendrá el tipo de valor '*character string*'. Los tipos de valor '*URI*', '*URN*' y '*URL*' también pueden ser de utilidad. Por ejemplo, el atributo '*description*' con el tipo de valor '*character string*' se puede utilizar para describir un término y la propiedad con el atributo '*extra information*' y el tipo de valor '*URL*' para indicar una página web con información sobre una categoría determinada.

5.4 Recapitulación y observaciones

En general, este apartado demuestra la utilidad, para la Terminología, de las herramientas informáticas orientadas a la representación de conocimiento. Observamos que las distintas propuestas han servido para poner en práctica diferentes aproximaciones a la representación de conceptos, por ejemplo, el enfoque teórico de Temmerman se refleja en la herramienta *ad hoc* para sus proyectos e incluye todos los elementos necesarios para reflejar un sistema conceptual desde ese punto de vista. Esto mismo ocurre con las otras dos herramientas analizadas. Esta exclusividad en la forma en la que se han desarrollado las herramientas supone, sin embargo, un obstáculo en cuanto a sus posibilidades de reutilización con otros enfoques. Otra desventaja es la poca facilidad de acceso a este tipo de software. Por ejemplo, en el caso de CODE4, la tecnología del programa no permite hacer uso de él en los sistemas modernos.

Independientemente de las posibilidades de uso de este tipo de herramientas, lo que cabe constatar es que este tipo de sistemas resultan de interés para la Terminología y suponen una nueva forma de representar conocimiento que se aleja de sistemas más tradicionales.

6 Ingeniería del conocimiento: las ontologías

La Ingeniería del Conocimiento (IC) es la disciplina tecnológica que se centra en una aproximación sistemática, disciplinada y cuantificable al desarrollo, funcionamiento y mantenimiento de sistemas basados en conocimiento (Alonso Betanzos et al., 2004: 11). Para ello, se trata de establecer metodologías que permitan abordar el desarrollo de sistemas basados en conocimiento de una forma más sistemática. El desarrollo de ontologías es uno de los sistemas que se abordan en esta disciplina. A continuación, exploramos sus antecedentes teóricos y las posibilidades de este tipo de sistemas en la representación de conceptos, dedicando un apartado específico a hablar de las ontologías.

6.1 El concepto de ‘característica’ en Ingeniería del Conocimiento

Antes de iniciar este bloque, conviene que revisemos el significado de *característica* desde el punto de vista de la Ingeniería del Conocimiento y otras denominaciones cercanas. De ahora en adelante, hablaremos de *propiedades, características o atributos* y conviene que sepamos diferenciar a qué conceptos hacen referencia estas denominaciones ya que, como hemos estudiado en Terminología, el concepto de una denominación desde un campo específico no siempre se corresponde con el de otro campo.

6.1.1 Cualidades y propiedades

Autores como Oltramari et al. (Oltramari et al., 2002: 20), influidos por vertientes filosóficas, distinguen entre *cualidad* y *propiedad* para hablar de las características conceptuales. Desde su punto de vista, la *cualidad* hace referencia a algo particular y la *propiedad* hace referencia a algo universal. Una entidad tiene ciertas cualidades, que existen mientras existe la entidad. Estas cualidades pertenecen a distintos *tipos de cualidad* (como color, tamaño, olor, etc.) y constituyen individuos específicos característicos de modo que dos entidades particulares no tienen una misma cualidad. Se distingue además entre una cualidad (por ejemplo, el color de una rosa concreta) y su valor (por ejemplo, un tono particular de rojo). El valor se denomina *quale* y describe la extensión (o clasificación) de la cualidad individual con respecto a un espacio conceptual particular (*espacio de cualidad*). Cuando dos rosas tienen el mismo color, se entiende que comparten el mismo *espacio de cualidad* pero aun así los autores distinguen dos cualidades distintas numéricamente. Es decir, comparten un mismo espacio de color aunque su cualidad de color es distinta, cada rosa tiene la suya (Masolo et al., 2003).

6.1.2 Cualidades y *qualia*

La distinción entre cualidades y *qualia* (*quale*) se inspira en la teoría de tropos monocategorial (Goodman, 1951; Campbell, 1990) y se justifica en su observación del lenguaje natural, que parece hacer una distinción similar a la de esta teoría. Por ejemplo, cuando se dice «el color de la rosa pasó de rojo a marrón en una semana» no se habla de un tipo de tono de rojo sino de algo que cambia su propiedad en el tiempo mientras sigue conservado su identidad. De este modo, los autores consideran que las cualidades son continuantes (*endurants*), están presentes en cualquier momento y no cambian (Masolo et al., 2003: 10-17). Mientras que cuando se dice «el rojo es distinto del verde» o «el rojo se parece al marrón», no se habla de cualidades sino de regiones dentro del espacio de una cualidad. Un tono particular de rojo que se observa como un átomo dentro del espacio de color.

Además, cada valor de una cualidad ocupará un región espacial de dicha cualidad (es decir, el valor 'rojo' estará situado cerca de 'marrón' y lejos de 'verde' dentro del espacio de 'color', siendo 'color' una cualidad). Por tanto, cada cualidad tendrá su región.

6.1.3 Propiedades y características

Nilsson (Nilsson, 2005), sigue esta misma línea y en su propuesta de modelización formal de un sistema de conceptos apunta hacia la importancia de distinguir entre clases, propiedades y características a la hora de definir formalmente cada sistema. Basa su planteamiento en teorías metafísicas y señala que las características pueden verse desde dos enfoques. Por un lado, está la teoría del sustrato, que entiende las características como elementos adscritos a algo existente anteriormente; por otro, la teoría de la agrupación, que entiende que la suma de características o tropos es lo que constituye una entidad individual. Desde el segundo enfoque, Nilsson entiende que se debe distinguir entre características que pertenecen a entidades individuales y propiedades que se atribuyen a universales. Nilsson intenta establecer una distinción entre las clases que constituirán una ontología acerca de un ámbito y las propiedades que servirán para describirlas. Para ello, analiza las diferencias desde distintos enfoques entre estos dos elementos. Desde un punto de vista lingüístico, las clases corresponden a nombres comunes, incluidos los verbos que expresan actividades o estados con entidades individuales o instancias designadas por nombres propios. Asimismo dirá que las propiedades básicamente se corresponden con adjetivos aunque a veces encontramos oraciones preposicionales que expresan la atribución de una propiedad. Según esto, prototípicamente un adjetivo + un nombre común identificarán una subclase. Sigue su trabajo explicando que desde el punto de vista de la lógica descriptiva tanto las clases como las propiedades se expresan como predicados unitarios (de un argumento) y por tanto, no hay ninguna diferencia a priori entre los dos elementos. Un predicado unitario de este tipo identifica un subconjunto determinado, sin importar si representa a una clase o a una propiedad. El predicado "azul" representa todas las cosas azules del universo modelo sin importar cuales sean. Sin embargo, desde el punto de vista de la modelización de datos y la programación orientada a objetos, es decir, con una base puramente orientada a la representación de conceptos mediante un lenguaje formal de programación sí existe una diferenciación visible entre clases y propiedades. Las propiedades serán *slots* que se heredan, puesto que las propiedades de una clase se heredan. Como conclusión, propone que las propiedades sirvan para clasificar subclases y que se hable de propiedades contingentes y propiedades esenciales. Las propiedades esenciales serán invariables mientras que las contingentes pueden variar según la situación del concepto en un sistema de predicados y pone de ejemplo *carbono*, una propiedad esencial será su número de átomos pero una contingente su color ya que puede ser negro o transparente según sea grafito o diamante.

Además, Nilsson (ibíd.) en su enfoque hacia la modelización de datos y partiendo de las matemáticas computacionales trata el espacio conceptual. El *espacio de un concepto* (Gärdenfors, 2004) es un espacio abstracto que debe visualizarse como un *espacio Euclidiano* en 3 dimensiones aunque tiene un dimensión que refleja las propiedades disponibles. Según esto, un concepto corresponde a una región de una noción y un subconcepto de un concepto se corresponde con una subregión coherente de la región del concepto. Visto así, las clases de distinguen de las propiedades porque tienen más de una dimensión y las propiedades conforman una sola dimensión.

6.1.4 Recapitulación y observaciones

Los trabajos incluidos en este apartado son una prueba de la cercanía existente entre los enfoques teóricos que parten de la Ingeniería del Conocimiento y los de la Terminología. Concretamente, en lo que se refiere al estudio de las características.

Como hemos mencionado en la introducción, la Ingeniería del Conocimiento es la disciplina que se ocupa de representar y organizar el conocimiento mediante sistemas informáticos y ello explica que la aproximación a las características se realice principalmente desde un punto de vista representacional. Anteriormente, hemos visto como Temmerman justificaba su planteamiento de distinguir entre nombres de característica y valores en la influencia de la Ingeniería del Conocimiento. Las similitudes entre enfoques terminológicos y de la Ingeniería del Conocimiento continúan a lo largo de todo el trabajo de Nilsson, primero respecto al doble rol de la característica, en sus palabras la distinción entre las características como elementos adscritos a algo que existe y las características como la suma de aquello que constituye una entidad individual. Nilsson también trata la dimensión lingüística y va más allá para debatir la correspondencia de una característica con un adjetivo, algo que también nos resulta familiar.

Otra similitud la encontramos en el lugar en el que se sitúan las características cuando se trata de representar conocimiento mediante un lenguaje formal. En este punto, las características tendrán el papel de acompañar a la entidad, en forma de *slot*, como un elemento agregado a una entidad que aporta información sobre su significado. Como cuando hacemos una definición de un término y utilizamos una suma de características para describirlo.

Finalmente, en el trabajo de Nilsson observamos de nuevo una similitud entre el espacio conceptual y la descripción de multi-dimensionalidad de un concepto introducida anteriormente (§1.2.5).

6.2 Representación de conceptos en Ingeniería del Conocimiento

Para hablar de representación de conceptos, necesitamos entender qué se entiende por ‘conocimiento’ en IC. Para ello, debemos distinguir entre dos puntos de vista tal y como lo expresa Guarino. Tradicionalmente el concepto de ‘conocimiento’ se había considerado desde una perspectiva funcional (Guarino, 1995: 625) donde la base de conocimiento se establecía de acuerdo no tanto con lo que era considerado cierto sino en lo que resultaba útil con respecto al objetivo del agente. Más tarde, autores como Clancey (Clancey, 1993) o Gruber (Gruber, 1995) hablan de Ingeniería del Conocimiento desde una perspectiva de modelado, donde la base de conocimiento no se considera un repositorio de conocimiento extraído de la mente de un experto sino el resultado de una actividad de modelización cuyo objetivo es el comportamiento observado por parte de un agente inteligente vinculado a un entorno externo. Es decir, la base de conocimiento tendrá que ser una realidad objetiva y no la mente del agente que lo expresa. En este caso, el conocimiento está más vinculado a una verdad clásica y no dependerá de la forma particular en que un agente entienda el dominio. En concreto, Guarino (ibíd.) señala que este tipo de modelado establece una correspondencia entre la base de conocimiento y dos subsistemas separados: el comportamiento del agente y su entorno. En este punto, el autor señala que algunas metodologías tienden a centrarse en el comportamiento del agente y dejar de lado la estructura inherente del dominio. Un beneficio potencial de estas metodologías es que solo la información relevante de un determinado dominio se tendrá en cuenta, pero como señala el autor, esto no tiene sentido si al menos una parte de la realidad objetiva representa aunque sea

parcialmente la estructura inherente del dominio. De no ser así, nos encontraríamos con que este planteamiento habría derivado justamente en aquello que se pretendía evitar: el modelado a partir del pensamiento de un agente.

Este planteamiento tiene la ventaja de que al representar el conocimiento teniendo en cuenta lo ‘verdadero’ (de acuerdo a la concepción aristotélica clásica de ‘verdadero’) se facilita la interacción y comunicación entre los distintos agentes que forman parte de las distintas comunidades dentro de un mismo dominio.

Además de ‘conocimiento’, en Ingeniería del Conocimiento, se habla de ‘datos’ y de ‘información’. Los *datos* son las señales sin interpretar que pueden llegar directamente a nuestros sentidos y la *información* se considera un dato provisto de significado. Un ejemplo aplicado en palabras de Alonso et al. (Alonso et al., 2004: 18) es el siguiente:

	característica	ejemplo
Dato	sin interpretar	luz roja de un semáforo
Información	añade significado a los datos	luz roja significa parar
Conocimiento	añade propósito y capacidad a la información potencial para generar acciones	si vamos conduciendo y nos encontramos un semáforo con luz roja, ejecutar las acciones para detener el vehículo si la luz roja está encendida, las demás luces deben estar apagadas

Tabla 4. Diferencia entre ‘dato’, ‘información’ y ‘conocimiento’

El conocimiento, por tanto, se interpreta como la capacidad de crear *información* a partir de la información interpretada.

Desde un punto de vista técnico, la representación de este conocimiento, se realiza mediante lo que se denomina *sistemas basados en conocimiento*. Se trata de sistemas software que sirven para conservar una gran cantidad de conocimiento y permiten explotar este conocimiento mediante métodos informáticos. Siguiendo el planteamiento teórico no funcional, un sistema basado en conocimiento consta de dos partes principales: la base de conocimiento y el motor de inferencia o proceso de razonamiento. Además, suele incluir otros componentes importantes como la base de hechos, la interfaz y el módulo de explicaciones. La base de conocimiento contiene tanto conocimiento factual como heurístico. El conocimiento factual es el conocimiento de un dominio aceptado por los expertos de ese campo. El conocimiento heurístico está basado en la experiencia y es menos riguroso ya que se extrae de la práctica o el razonamiento plausible (ibídem: 14).

A nivel metodológico, el conocimiento adquirido sobre una materia debe ser organizado y preparado para que pueda ser utilizado computacionalmente en un sistema. En Ingeniería del Conocimiento, cuando se desarrolla un Sistema Basado en Conocimiento, se intenta ver el problema desde el punto de vista del usuario y después desde el punto de vista del sistema. Para llevar a cabo esta labor se realiza una modelización que se estructura en dos partes: el modelado conceptual y el modelado formal.

El modelado conceptual es una representación del conocimiento del dominio en estructuras no computables que modelan su conducta en la solución de problemas. En una segunda fase, se realiza la formalización que consiste en transformar los modelos conceptuales

en modelos formales. Formalizar consiste en representar simbólicamente el conocimiento mediante alguno de los formalismos existentes, organizarlo de acuerdo con algún modelo de diseño y determinar los métodos de inferencia adecuados para manejar eficientemente dichos conocimientos.

El modelado formal es, por tanto, una representación semi-computable del conocimiento. Para que un modelo formal sea totalmente operativo es necesario un modelo computable, formado por bases de conocimiento, motores de inferencia y estrategias de control. Una vez se ha realizado el proceso de adquisición del conocimiento, a través de expertos u otras fuentes y se ha creado el modelo conceptual que representa los componentes del dominio creado, se debe buscar un modelo semi-computable que tenga la suficiente capacidad expresiva para contener el conocimiento necesitado por el sistema y que permita realizar procesos de inferencia para deducir nueva información o resolver el problema que tiene como meta la creación del sistema (ibídem: 21). Para ello se utilizan, por ejemplo, los esquemas de representación, entendidos como paradigmas que codifican la realidad en un ordenador.

6.2.1 Esquemas de representación formal

En los *esquemas lógicos*, los hechos se presentan como fórmulas lógicas en algún formalismo matemático. En este caso la base de conocimiento será una colección de fórmulas lógicas bien formadas y verdades que describen un dominio (ibídem: 23).

En Ingeniería del Conocimiento se trabaja con distintos esquemas de representación que provienen de la Inteligencia Artificial como, por ejemplo, los esquemas lógicos, las redes semánticas o los marcos. Para escoger entre uno u otro, los ingenieros deben considerar el objetivo del sistema, el tipo de conocimiento y la eficiencia del esquema de representación (Rich y Knight, 1991). A continuación, describimos brevemente algunos de los más utilizados para centrarnos en el próximo apartado en la lógica de descripciones, de modo que podamos entender más adelante las bases de una ontología informática.

6.2.2 Lógica de primer orden

La lógica de primer orden es uno de los esquemas de representación más utilizados ya que es intuitivamente entendible y usa pocos elementos sintácticos. Los métodos de resolución son simples y permiten realizar un seguimiento de los pasos de resolución. La incorporación o eliminación de nuevo conocimiento es muy sencilla, basta agregar o eliminar el hecho concreto en la base de conocimiento. El procedimiento para resolver inferencias es la resolución por refutación, el procedimiento trabaja con fórmulas representadas de forma *clausal* (disyunciones o negaciones). La desventaja de la lógica de primer orden es que en la resolución de determinados problemas o razonamientos ésta resulta ser indecidible (Hayes, 1979), es decir, que no se algunos problemas no se pueden resolver mediante la lógica de primer orden.

6.2.3 Redes semánticas

Las *redes semánticas* o *grafos conceptuales*, atribuidas a Quillian (Quillian, 1967), se basan en la idea de que el significado de un concepto depende de las relaciones que tiene con otros de la base de conocimiento. Se diseñaron como modelos psicológicos de la memoria humana con el fin de dotar de semántica a las palabras, muchos de los trabajos en esta línea van encaminados al procesamiento del lenguaje natural.

La base de representación consiste en modelar los conocimientos relativos a un objeto o concepto mediante pares atributo-valor. Los pares se representan en el grafo de la siguiente forma: el nodo origen es el objeto o concepto para el que se definen los pares atributo-valor (p. ej. *película_clásica*); los arcos que parte de dicho nodo son los atributos del par (p. ej. *Filmada_en*); y los nodos destino representan los valores de los atributos (p. ej. *Cinemascope*). Los arcos se agrupan en dos categorías:

- arcos descriptivos: describen conceptos mediante atributos propios del dominio (*rodada_En, año_Rodaje*)
- arcos estructurales: enlazan los conceptos formando la estructura de la red. La semántica de arcos estructurales es independiente del dominio que se está representado. Las etiquetas más comunes son:
 - instanciación: el arco *Instancia* liga un objeto concreto con su tipo o clase genérica
 - generalización: relaciona una clase con otra más general, la relación más utilizada es *Subclase-de*
 - agregación: enlaza un concepto con sus componentes. La relación más común es *parte-de*

Además de las representaciones de relaciones binarias que permiten formar una taxonomía con los conceptos de dominio mediante las relaciones estructurales y expresar las características con relaciones descriptivas, las redes semánticas permiten representar acciones o frases complejas. Esto se basa en la teoría de que una frase tiene una estructura formada por un verbo relacionado con uno o más sintagmas nominales, denominados *casos*. Esta forma de representación se denomina *gramática por casos* y se utiliza en procesamiento del lenguaje natural. Los casos definidos van ligados a un verbo y pueden ser, por ejemplo, *agente, contra-agente, objeto, resultado, instrumento, origen, propósito, lugar, tiempo, sujeto, modalidad*.

El uso de las relaciones estructurales para formar jerarquías de conceptos permite acceder a propiedades en otros nodos más generales utilizando los arcos *instancia* y *subclase-de*. Así, evitamos repetir propiedades en la base de conocimiento, a este fenómeno se le denomina *herencia de propiedades*.

El valor de una propiedad definida en un concepto es válido, por regla general, para todos los conceptos e instancias situados en niveles inferiores del árbol taxonómico. Para representar la excepción, en una instancia o concepto se redefine la propiedad con el valor propio que contenga ese elemento.

La forma que tienen las redes semánticas de responder a consultas sobre el conocimiento almacenado es mediante la técnica de la equiparación. Para realizar la equiparación se crea un fragmento de red semántica con los valores conocidos y no conocidos de la pregunta y se intenta emparejar con alguna sección de la red semántica que encaje.

Este tipo de esquemas se denominan *declarativos* porque el conocimiento sobre el dominio se representa como un conjunto estático de hechos y se separa de los métodos o procedimientos utilizados para realizar inferencias sobre estos hechos.

6.2.4 Marcos

Los marcos o *frames*, cuyo máximo precursor fue Minsky (Minsky, 1975), es la técnica de representación más utilizada en Inteligencia Artificial. Poseen una gran capacidad para representar aspectos semánticos sobre el conocimiento *declarativo* del dominio y además, permiten representar conocimiento *procedimental* que se refiere a cómo rellenar los valores de los atributos, cómo realizar comprobaciones sobre la información introducida, qué acciones realizar cuando se cumplan o fallen los procesos de búsqueda, etc.

Los elementos que aparecen en esta técnica son:

- *marcos*, que representan conceptos o instancias de conceptos
- *relaciones*, expresan dependencias entre conceptos
- *propiedades*, atributos de conceptos
- *facetas*, representan restricciones sobre los valores que tomarán las propiedades

En cuanto a los marcos, existen dos tipos:

- *marcos clase*, representan conceptos, clases o situaciones genéricas, se describen mediante un conjunto de propiedades (que pueden tener o no valores asignados) comunes al concepto que representan.
- *marcos instancia*, representan elementos específicos de los marcos clase y sus propiedades se rellenan con información concreta de los individuos a los que representan.

Las relaciones se definen entre marcos clase y/o marcos instancia. Se entienden como propiedades especiales cuyo valor es otro marco. Hay relaciones estructurales que son independientes del dominio como en las redes semánticas pero también hay relaciones más específicas del dominio representado. Las relaciones estructurales son fijas y están definidas en el sistema de representación, *subclase-de* o *instancia-de* son ejemplos. En estos sistemas podemos encontrar también relaciones inversas entre ellas como *superclase-de* para *subclase-de*.

Para realizar la taxonomía se suelen utilizar las relaciones *subclase-de* que permiten representar el dominio mediante la especialización de conceptos generales en otros más específicos. En esta relación, los marcos origen y destino deben ser *marcos clase*. Igual que en las redes semánticas, estas relaciones definen el camino de la herencia de propiedades.

La relación *instancia-de* se establece entre un marco instancia y un marco clase, indicando que el marco instancia es un elemento del conjunto o clase representado en el marco clase. Un marco instancia puede serlo de varios marcos clase. Además de éstas, existen otros tipos de relaciones estructurales que representan dependencias o vínculos organizativos entre conceptos de un dominio determinado. En la mayoría de herramientas, estas relaciones no suelen estar definidas, por lo que los motores de inferencia no trabajan con ellas directamente, para ello, el ingeniero tendrá que diseñar las inferencias con procedimientos propios, reglas o dominios, o con alguna forma de expresar los modos procedimentales. Según Alonso et al. (Alonso et al., 2004: 38), ejemplos de estas relaciones son:

- *Relación disjunta*: definida entre marcos clase, indica que no pueden tener marcos instancia comunes.

- *Relación exhaustivo*: definen completamente a la superclase, por lo que no puede existir una instancia directa de la superclase.

También existen otras relaciones, denominadas *ad hoc*, que expresan dependencias a medida entre conceptos de un dominio. Estas relaciones se definen entre marcos clase y se instancian en los marcos instancia.

Las *propiedades* (slots) o *atributos* se definen en los marcos clase y caracterizan tanto a los marcos clase como a los marcos instancia. Pueden ser de dos tipos:

- Propiedades de clase: representan atributos o características genéricas de un concepto o clase. Se definen y rellenan en el marco clase y tendrán el mismo valor en todas las instancias de esa clase.
- Propiedades de instancias: Se definen en el marco clase pero se rellenan en los marcos instancia con valores concretos. No es necesario rellenas todas las propiedades de un marco instancia si no se han definido como obligatorias.

6.2.5 La lógica de descripciones

La lógica de descripciones (lógica descriptiva, *description logics*, *DL*) surge de la necesidad de identificar un fragmento de la lógica de primer orden lo suficientemente expresivo y con las propiedades computacionales necesarias para representar y razonar sobre conjuntos de conceptos, sus relaciones y las instancias o clases concretas (Baader et al., 2003; Jiménez-Ruiz, 2010).

Jiménez-Ruiz además nos recuerda que originalmente la Lógica de descripciones se denominó lenguaje terminológico o lenguaje conceptual (Jiménez-Ruiz, 2010).

La DL provee teorías y sistemas para expresar información estructurada, permitir su acceso, y poder razonar de forma semánticamente precisa. Las representaciones en DL serán a nivel de predicados, sin variables ni funciones. Cada vez que se defina un formalismo lógico nuevo existirá una compensación entre mayor poder expresivo y menor coste computacional, lo que permite una gran eficiencia de razonamiento.

La DL destaca principalmente por su semántica basada en una lógica bien, con unos límites claros y decidibles. Esto permite inferir conocimiento representado de forma implícita en a partir de conocimiento representado explícitamente en DL, mediante el uso de razonadores. La DL soporta el uso de patrones de inferencia, un elemento muy utilizado por las aplicaciones de los sistemas de procesamiento de información y también por los humanos cuando queremos estructurar y entender el mundo clasificando sus conceptos e individuos y buscando patrones que nos permitan detectar información implícita acerca de ellos.

En DL, la clasificación de conceptos determina relaciones sub-conceptos – superconcepto entre conceptos de una determinada terminología, lo que permite estructurar la terminología en forma de jerarquía de subsunción.

Asimismo, la clasificación de individuos, también denominados *objetos*, determina si un individuo es siempre instancia de un determinado concepto, la relación de instancia se deduce por la descripción del individuo y del concepto. Esto proporciona información útil sobre las propiedades de un individuo y además las relaciones de las instancias de un concepto pueden activar la aplicación de reglas que aportan información adicional a la base de conocimiento.

Con el uso de la DL se pretenden superar las ambigüedades que surgían con el uso de redes semánticas y de marcos tradicionales, para ello, DL basa su desarrollo en tres ideas surgidas del trabajo de Brachman y Schmolz (Brachman y Schmolz, 1985: 171-126), que estudiaban el comportamiento de las redes con herencia estructurada.

Los bloques de construcción sintáctica básicos que son también tres:

- Conceptos atómicos (predicados unitarios, nombres de conceptos)
- Roles atómicos (predicados binarios, nombres que hacen referencia a roles)
- Individuos (entendidos como las constantes)

El objetivo es restringir el poder expresivo del lenguaje mediante un conjunto pequeño de constructores (adecuados epistemológicamente) para construir conceptos y roles.

Y finalmente, que el conocimiento implícito sobre los conceptos y los individuos se pueda inferir automáticamente con la ayuda de procedimientos de inferencia. En concreto, las relaciones de subsunción entre conceptos y las relaciones de instancia entre individuos y conceptos juegan un papel muy importante. A diferencia de los enlaces tipo *IS-A* (es un) introducidos explícitamente por las redes semánticas, las relaciones de subsunción y las relaciones de instancia se pueden inferir de la definición de conceptos y las propiedades de individuos.

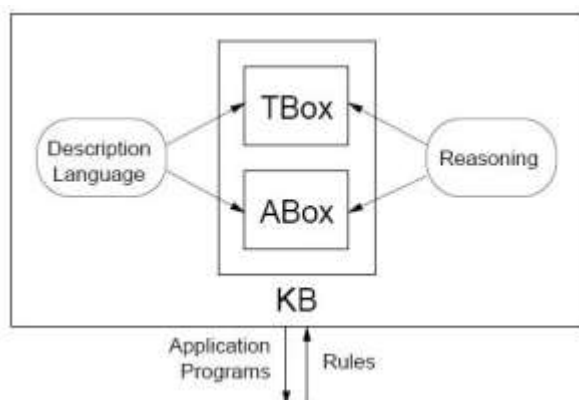


Figura 15. Arquitectura de un sistema de representación basado en la lógica de descripciones (Baader y Nutt, 2003: 50)

La arquitectura de un sistema de representación basado en DL (figura 15) se compone por una base de conocimiento que a su vez se divide en dos elementos: uno que se denomina *TBox* y se encargará de almacenar la terminología (conocimiento intensional), y otro denominado *ABox*, donde se encuentran las aserciones sobre los individuos mencionados en la terminología (conocimiento extensional).

TBox contiene tanto conceptos que denotan conjuntos de individuos como los roles que denotan las relaciones binarias entre individuos.

La DL permite construir descripciones complejas ya que las declaraciones en *TBox* y en *ABox* pueden identificarse con fórmulas de la LPO. El razonamiento en DL se utiliza normalmente para determinar si una descripción es adecuada (no contradictoria), si una descripción es más general que otra o si subsume a otra. Los problemas que podemos encontrar en este tipo de sistemas y sobre todo en el módulo del *ABox* se deben al uso de afirmaciones

inconsistentes. Para ello, se pueden utilizar reglas que funcionen como una extensión del núcleo lógico y restringirán las descripciones.

6.2.6 Recapitulación y observaciones

En este apartado hemos visto, a grandes rasgos, los elementos que configuran un sistema de bases de conocimiento. Entendemos estos sistemas como un producto informático que tiene información organizada de un modo que permite que un ordenador pueda realizar tareas con ella, como por ejemplo, recuperar información, estructurar datos, inferir información, etc. Estos sistemas se componen de una base de conocimiento y de un motor de inferencia. La base de conocimiento contiene conocimiento heurístico, reglas y algoritmos, y conocimiento factual, conocimiento de un dominio. No cabe duda de la utilidad de estos sistemas para el trabajo terminológico pero además, desde nuestro punto de vista, el terminólogo tiene un papel importante en el desarrollo de estos sistemas, por su capacidad y formación en organización y descripción de conceptos.

A lo largo de este apartado, nos hemos familiarizado con conceptos como la formalización, representar de forma que una máquina pueda interpretar. Finalmente, hemos visto distintos ejemplos de esquemas de representación formal de conocimiento y de forma más detallada la lógica de descripciones. De las descripciones presentadas, podemos observar que existen distintos modos de formalizar el conocimiento y dependerá de la finalidad de la formalización, escoger uno u otro.

La utilización de la lógica de descripciones tiene como objetivo expresar nuestro conocimiento de una forma restringida y formal. Su principal ventaja es que permite la clasificación de conceptos de un modo similar al que utilizamos en Terminología cuando estructuramos conceptos en un tesoro y además, permite inferir conocimiento representado de forma implícita a partir de conocimiento representado explícitamente. Es decir, podemos representar en forma de tesoro los conceptos de un dominio e inferir a partir de relaciones que hayamos establecido en el “tesoro”, nuevas relaciones entre conceptos del “tesoro” que se deduzcan de estas.

6.3 Ontologías

Así como en Terminología tradicionalmente se han elaborado bases de datos o diccionarios para estructurar y organizar nuestros términos; en Ingeniería del Conocimiento se hace uso de sistemas como las ontologías y, por ende, de la ingeniería ontológica, para representar el conocimiento. La Ingeniería Ontológica hace referencia al conjunto de actividades implicadas en el proceso de desarrollo de una ontología, su ciclo de vida y las metodologías, herramientas y lenguajes para construir ontologías (Gómez-Pérez, Fernández-López et al., 2004:5), y es una rama de la Ingeniería del Conocimiento.

El término ‘ontología’ proviene de la Filosofía y se refiere a una explicación sistemática de lo que existe, pero durante las últimas décadas, esta palabra ha tomado relevancia para la Ingeniería del Conocimiento aunque, a menudo, alude a distintos conceptos. Guarino y Giaretta (Guarino y Giaretta, 1995: 25) identifican siete significados distintos para la palabra “ontología”:

- 1. Ontology as a philosophical discipline*
- 2. Ontology as an informal conceptual system*

3. *Ontology as a formal semantic account*
4. *Ontology as a specification of a conceptualization*
5. *Ontology as a representation of a conceptual system via a logical theory*
 - a. *characterized by specific formal properties*
 - b. *characterized only by its specific purposes*
6. *Ontology as the vocabulary used by a logical theory*
7. *Ontology as a (meta-level) specification of a logical theory*

La primera acepción es de la Filosofía mientras que las demás (2-7) aluden a la Inteligencia Artificial, la Ingeniería del Conocimiento o la Web Semántica, entre otras disciplinas informáticas. Gómez-Pérez et al. (Gómez-Pérez et al., 2004:5) explican que éstas (2-7) y otras definiciones aparecidas en la literatura, se refieren a una misma realidad observada desde distintos puntos de vista y con matices que la complementan.

Una de las definiciones más citadas cuando se define ontología en este campo es la de Gruber (Gruber, 1993: 199): «*An ontology is an explicit specification of a conceptualization*». Esta definición ha sido revisada por Borst (Borst, 1997: 12) que añade que las ontologías son una especificación formal de una conceptualización compartida. Más tarde, Studer et al. (Studer et al., 1998: 185) basándose en estas dos definiciones describen una ontología como: «*a formal, explicit specification of a shared conceptualization*». Y continúan su definición explicando que una conceptualización (*conceptualization*) es un modelo abstracto de un fenómeno del mundo del cual se han identificado sus conceptos relevantes. Explican que es explícito (*explicit*) porque alude a que el tipo de conceptos utilizado y las restricciones de su uso están definidos explícitamente. Formal (*formal*), porque apunta al hecho de que una ontología debe ser procesable por un ordenador. Y al decir conceptualización compartida (*shared*) persiguen reflejar la idea de que una ontología captura un conocimiento consensuado, no privado o individual sino aceptado por un grupo.

Según Gómez-Pérez et al. (Gómez-Pérez et al., 2004: 1), las ontologías resultan útiles en los sistemas basados en conocimiento porque permiten modelar el conocimiento declarativo (objetos y hechos de un dominio). Las ontologías combinadas con sistemas de resolución de problemas, que permiten mecanismos de razonamiento específico, pueden configurar sistemas basados en conocimiento reutilizables.

Existen muchas tipologías para distinguir entre tipos de ontología y entre ellas, encontramos dos que clasifican las ontologías según la riqueza de su estructura interna y según el sujeto de la conceptualización (Gómez-Pérez et al., 2004: 28).

La riqueza de la estructura interna de las ontologías se refiere a su constitución y al nivel descriptivo de la ontología. Según la carga de información semántica y descriptiva se hablará de ontologías pesadas y ontologías ligeras (Lassila y McGuinness, 2001) y entre de estos dos polos encontramos (de más ligeras a más pesadas):

- Vocabularios controlados. Una lista finita de términos
- Glosarios
- Tesoros. Añaden información semántica pero carecen de jerarquías.

- Jerarquías no formales de tipo *IS-A* (es un). La relación “es-un” dependerá de la aplicación concreta de la ontología y no es estrictamente lógica desde un punto de vista conceptual. Por ejemplo, ‘alquiler de coches’ y ‘hotel’ pueden modelarse como tipos de viaje por razones prácticas.
- Jerarquías formales de tipo *IS-A* (es un).
- Marcos conceptuales. Contienen clases, propiedades, herencia y una taxonomía formal de tipo *IS-A*
- Ontologías que expresan restricciones de valores. Por ejemplo, en una propiedad: fecha de llegada siempre es una fecha
- Ontologías que expresan restricciones lógicas generales. Son las más expresivas e incluirán restricciones lógicas, según la lógica de primer orden.

En esta línea, el trabajo de Jimeno-Yepes et al. (Jimeno-Yepes et al., 2009) diferencia entre tipos de ontología según su nivel de expresividad. En la figura 16, se observa la clasificación de ontologías y recursos terminológicos del ámbito médico. A la izquierda se han situado aquellos recursos que contienen información terminológica o lingüística y tienen un menor grado de formalidad, considerando como menos formal al glosario. A la derecha se han situado los más formales, siendo el más formal la lógica compleja.

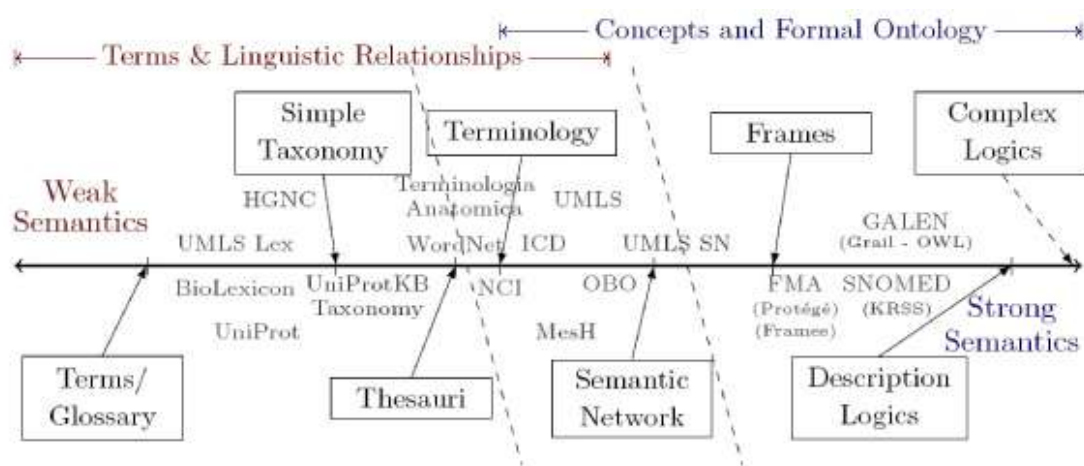


Figura 16. Espectro de ontologías basado de Jimeno-Yepes et al. (Jimeno-Yepes et al., 2008:2)

Si en cambio, centramos el foco de interés en el sujeto de la conceptualización, es decir, en el contenido informativo que se representará, encontramos:

- Ontologías de representación de conocimiento (van Heijst et al., 1997). Proporcionan definiciones formales para las primitivas de representación utilizadas en lenguajes basados en marcos, un ejemplo sería *Frame Ontology* (Gruber, 1995).
- Ontologías generales o comunes. Se utilizan para representar conocimiento del sentido común reutilizable entre dominios. Un ejemplo sería la *Meteorology Ontology* (Borst, 1997)
- Ontologías top-level o upper-level. Describen conceptos muy generales y ofrecen nociones bajo las cuales construir más conocimiento terminológico.
- Ontologías de dominio (Mizoguchi et al., 1995; van Heijst, Schreiber et al., 1997). Estas ontologías son reutilizables en su dominio específico,

proporcionan vocabulario sobre conceptos dentro de un dominio y sus relaciones, actividades, teorías y principios elementales dentro de ese dominio. Los conceptos de las ontologías de dominio son especializaciones de conceptos ya definidos en ontologías *top-level*, lo mismo ocurre con sus relaciones.

- Ontologías de tareas (Guarino, 1995; Mizoguchi et al., 1995). Describen una tarea o actividad concreta.
- Ontologías tarea-dominio. Describen una tarea específica de un dominio.
- Ontologías de métodos. Proporcionan definiciones de conceptos relevantes a un proceso de razonamiento concreto para realizar una tarea.
- Ontologías de aplicación (van Heijst et al., 1997). Dependen de una aplicación concreta y en principio solo servirán para ésta.

En el trabajo de Gómez-Pérez et al. (ibíd.) observamos además una tercera tipología de ontologías a lo largo de su apartado de ontologías más relevantes:

- ontologías de representación de conocimiento
- ontologías top-level
- ontologías de dominio
- ontologías lingüísticas

Esta forma de clasificar las ontologías se basa en su objetivo final de la ontología y, como vemos, diferencia entre representar conocimiento general, representar una realidad o un dominio y representar lenguaje.

6.3.1 Ejemplos de ontologías

Siguiendo la última tipología de ontologías, dentro de las **ontologías** consideradas **de representación de conocimiento**, Gómez-Pérez y otros agrupan a RDF (Manola y Miller, 2004) y a OWL (McGuinness y Van Harmelen, 2004). En realidad, OWL o RDF se consideran, generalmente, como lenguajes de representación de ontologías pero por eso mismo, se componen de un modelo de datos que constituye una ontología en sí misma. En otras palabras, estos modelos son más que ontologías de primer nivel porque ellos mismos constituyen el marco de trabajo, el lenguaje en el que se expresaran posteriormente otras ontologías. Por ejemplo, en el caso de OWL, se compone de los elementos que servirán posteriormente para agrupar conceptos en la ontología a través de clases, propiedades o relaciones o crear axiomas para describir estos elementos. Si observamos el ejemplo de taxonomía de OWL, nos podemos hacer una idea de que los elementos que aparecen son los instrumentos, las primitivas necesarias que nos ayudaran a formalizar una ontología computacionalmente (figura 16):

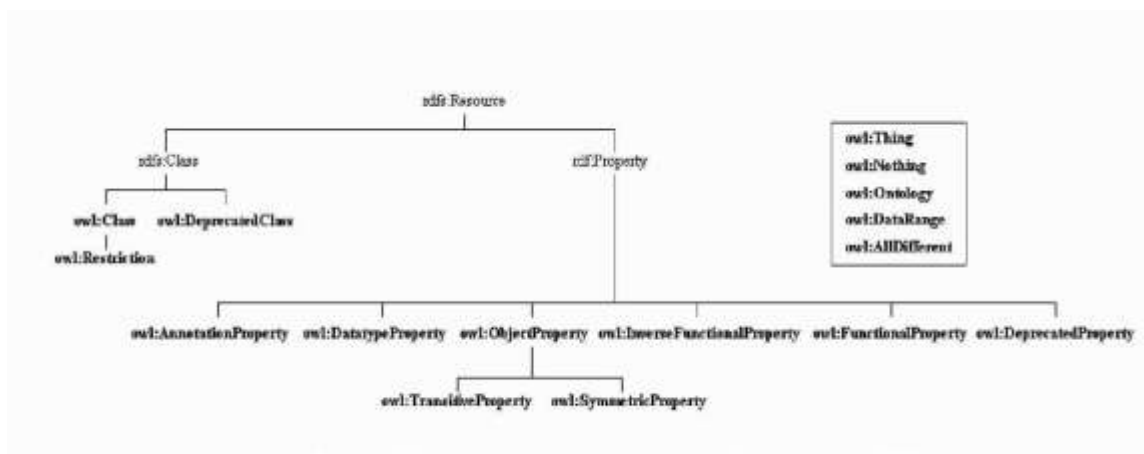


Figura 17. Taxonomía de clases de OWL-KR (Gómez-Pérez et al., 2004: 67)

Por otra parte, se considera **ontologías top-level** a aquellas que describen conceptos generales del mundo que están presentes en más de un dominio. A menudo, estas ontologías se utilizan como base para elaborar ontologías de dominio. Un ejemplo de ontología de primer nivel es la ontología de Sowa (Sowa, 1999: 67-109) que se muestra a continuación (figura 18).

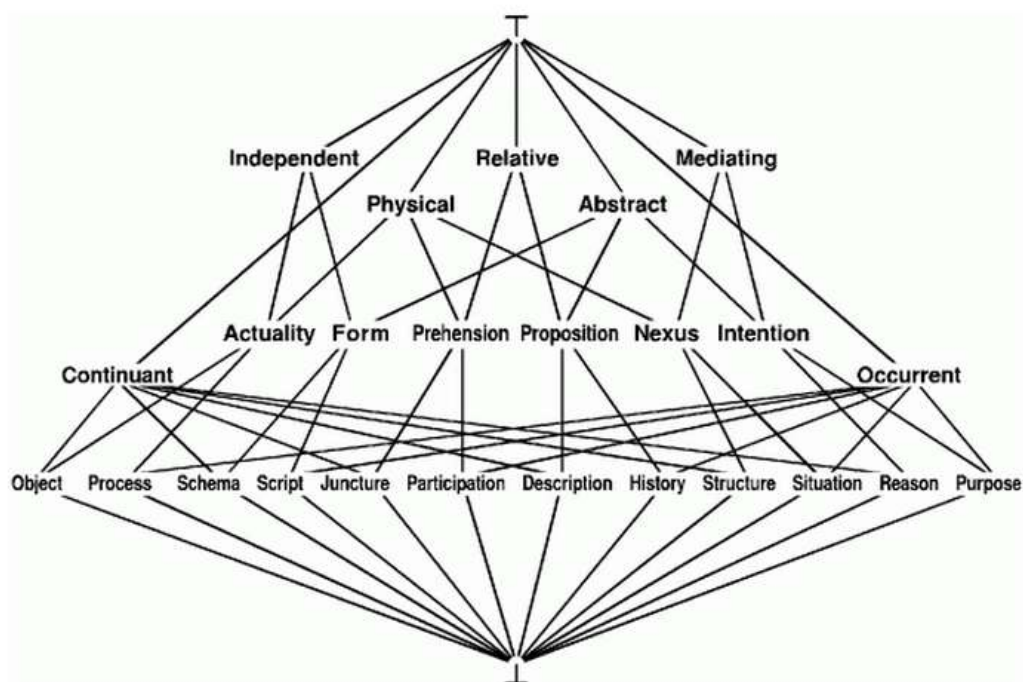


Figura 18. Extracto de la jerarquía de categorías top-level de Sowa

En el caso de las **ontologías de dominio**, los ejemplos son mucho más numerosos. Uno de los dominios con más proliferación de ontologías es el de la medicina. Las aplicaciones en este campo han demostrado ser de gran utilidad para la práctica médica y encontramos ontologías como UMLS (*Unified Medical Language System*). UMLS se compone de un meta-tesauro, una red semántica y un conjunto de léxico especializado (Bodenreider et al., 2004).

- El metatesauro contiene información biomédica sobre cada uno de los términos incluidos en UMLS. En 2011, UMLS contiene más de 2 millones de conceptos en su sistema a los cuales podemos acceder desde el meta-tesauro. Los conceptos en UMLS

se vinculan a denominaciones en distintos idiomas, lo que supone más de 10 millones de denominaciones entre los 21 idiomas en los que trabaja el sistema.

- La red semántica es una ontología, que se entiende como ontología de tipo *top-level* y se compone de conceptos biomédicos y de relaciones entre esos conceptos. En realidad, esta denominación se debe sobre todo a la finalidad de la ontología, ya que es la receptora de los conceptos presentes en el meta-tesauro y será en esta ontología donde se estructure y se describa la relación entre los conceptos del meta-tesauro. La ontología no se basa en ningún recurso textual sino que se creó ad hoc por especialistas del campo para poder estructurar la información del sistema.
- El lexicón especializado es un compendio de información sintáctica sobre términos biomédicos que tiene como finalidad ser aplicado en el procesamiento del lenguaje natural.

A continuación (figura 19), mostramos un extracto de la red semántica de UMLS, donde se puede observar la estructuración de las categorías de conceptos y como se distingue entre relaciones de hiperonimia y otras relaciones.

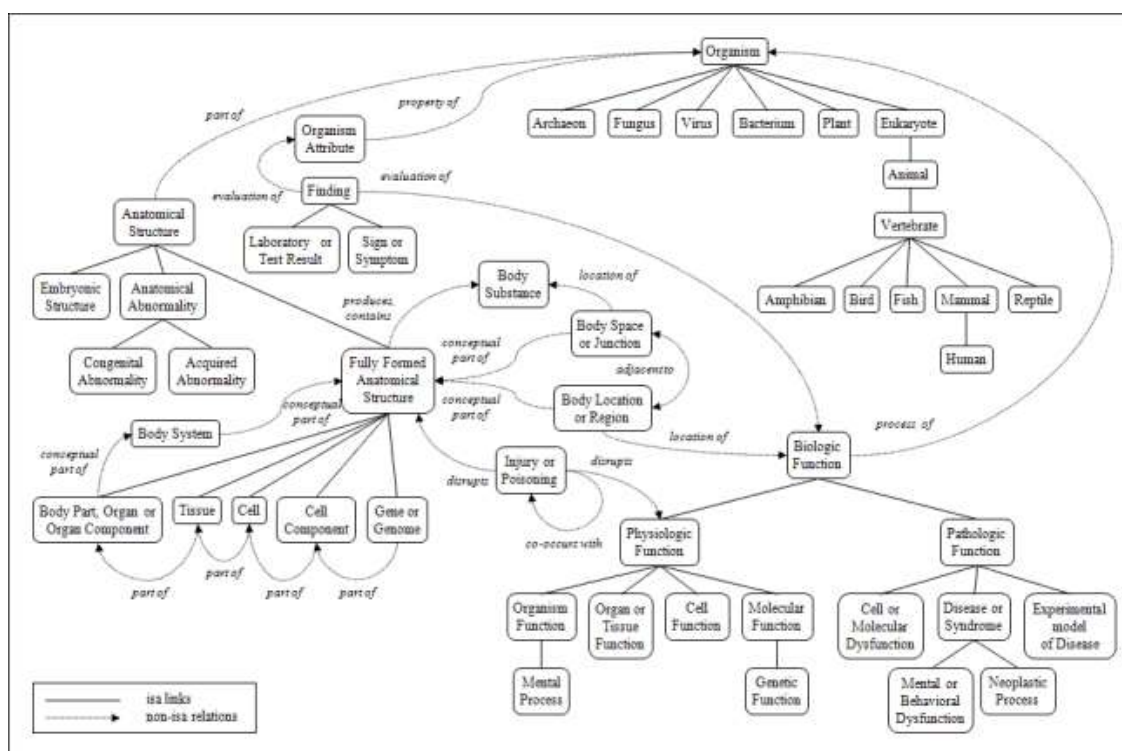


Figura 19. Representación de un extracto de la red semántica de UMLS (UMLS, 2009)

La diferencia entre este tipo de ontologías y las **ontologías lingüísticas** recae en que éstas últimas tienen como objetivo describir constructos semánticos y no modelar un dominio concreto. En estas ontologías no se formaliza el conocimiento de un dominio sino el modo en que este conocimiento se representa a través del lenguaje. Según Gómez-Pérez et al., la principal característica de las ontologías lingüísticas es que están ligadas a la semántica de las unidades gramaticales (palabras, sintagmas nominales, adjetivos, etc.).

La mayoría de ontologías lingüísticas utiliza la palabra como unidad gramatical, aunque hay casos de ontologías lingüísticas donde se utilizan unidades mayores que la palabra, como

GUM (Bateman et al., 1995: 60-72) o SENSUS (Swartout y Gil, 1995) que además se pueden considerar ontologías de tipo *top-level*, ya que contienen conceptos muy abstractos.

Este tipo de ontologías suelen basarse en el significado de las palabras y establecen un mapeado entre las denominaciones posibles y el concepto al que se refieren. Wordnet y Eurowordnet son ejemplos de ontologías de este tipo, y se distinguen entre ellas principalmente porque las ontologías como Wordnet dependen de una única lengua y, en cambio, Eurowordnet es multilingüe. Mikrokosmos también se podría agrupar en este tipo de ontologías lingüísticas, aunque en este caso, Mikrokosmos se caracteriza por ser independiente del idioma con que se trabaje.

Wordnet (Fellbaum, 1998) parte como un proyecto monolingüe en inglés que posteriormente se ha desarrollado también para otras lenguas. Empezó como una gran base de datos léxica en inglés desarrollada por la Universidad de Princeton y se fundamenta en teorías psicolingüísticas cuya finalidad es organizar los términos según su significado en lugar de según su forma.

Wordnet 3.0 en inglés contiene alrededor de 200.000 palabras agrupadas bajo alrededor de 117.000 *synsets* (conjuntos de sinónimos) que han servido para representar alrededor de 150.000 conceptos.

Cada *synset* está interrelacionado mediante relaciones de tipo sinonimia, antonimia, hiperonimia, hiponimia, meronimia u holonimia. Además, incluyen descripciones sobre su significado y están categorizadas en sustantivos, adjetivos, adverbios y lo que se denomina *palabras funcionales*. Cada categoría tiene un subconjunto de categorías predefinido bajo el cual se agruparán el resto de palabras. Por ejemplo, en el caso de los sustantivos en inglés, las categorías principales son las que aparecen a continuación (Miller, 1998: 29):

{act, action, activity}	{natural object}
{animal, fauna}	{natural phenomenon}
{artifact}	{person, human being}
{attribute, property}	{plant, flora}
{body, corpus}	{possession}
{cognition, knowledge}	{process}
{communication}	{quantity, amount}
{event, happening}	{relation}
{feeling, emotion}	{shape}
{food}	{state, condition}
{group, collection}	{substance}
{location, place}	{time}
{motive}	

En el caso de los adjetivos, las categorías principales bajo las que se agrupan el resto de adjetivos son adjetivos descriptivos y adjetivos relacionales. Una peculiaridad de este tipo de categoría es que además de agruparse en *synsets*, en el caso de representar la antonimia, es la categoría de las denominaciones la actúa como cabecera y se enlaza a la cabecera del *synset* antónimo, por ejemplo (figura 20):

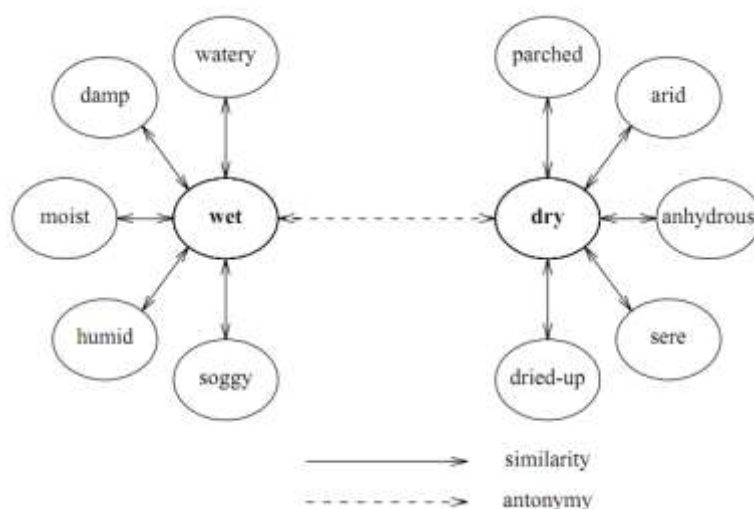


Figura 20. Estructura doble para un adjetivo en *Wordnet*

6.3.2 Recapitulación y observaciones

En este capítulo hemos analizado el concepto de ontología y hemos visto la variedad de definiciones y tipos que existen alrededor de la palabra *ontología*. Una de las conclusiones a las que llegamos es que no siempre el concepto de ontología abarca el mismo tipo de objeto, o dicho de otra manera, si nos fijamos en los ejemplos de ontologías que proponen Gómez-Pérez et al., algún autor más cercano a la Terminología que a la Ingeniería podría no coincidir al considerar *Wordnet* una ontología o acercarse a *OWL* como un ejemplo de ontología en lugar de como un mero lenguaje de formalización.

En lo que respecta a los distintos tipos de ontología presentados en este apartado, observamos cómo se reafirma la idea de que la representación del conocimiento de un dominio es un objetivo compartido por la Ingeniería Ontológica y la Terminología. Igualmente, consideramos que clasificaciones como la de Lassila y McGuinness o Jimeno-Yepes et al. nos permiten discernir las diferencias entre un recurso no formal y una ontología formal.

En esta línea, un aspecto que no podemos olvidar se expresa en la definición de ontología de Studer que completa la tradicional definición de Gruber: la ontología como la especificación formal de un conocimiento consensuado. Si observamos una ontología como la especificación formal de un conocimiento consensuado, podemos entender el papel de la Terminología a la hora de formalizar en palabras ese conocimiento consensuado.

Resulta muy interesante también la clasificación de Gómez-Pérez al estructurar su trabajo en ontologías de representación de conocimiento, top-level, lingüísticas y de dominio. En este aspecto parece coincidir con autores como Roche que abogan por diferenciar la representación del lenguaje de uso, de la realidad de un dominio o de la forma en que esta realidad se plasma en un sistema informático.

Desde nuestro punto de vista, una ontología terminológica debe inspirarse en las ontologías de dominio de la Ingeniería del Conocimiento de manera que nos permitirá un nivel de abstracción independiente del lenguaje y podamos dar prioridad a la representación de los conceptos que subyacen bajo los términos de un ámbito de especialidad. Al mismo tiempo, esta ontología ha de permitir conectar las denominaciones y otros aspectos lingüísticos y comunicativos con lo que se tendrá que mantener un equilibrio entre la consistencia formal, la naturaleza de los conceptos descritos en la ontología y el propósito de ésta.

7 El lenguaje estándar de ontologías OWL

El Consorcio *World Wide Web* (W3C) desarrolla tecnologías inter-operativas (especificaciones, líneas maestras, software y herramientas) para guiar la Red a su potencialidad máxima a modo de foro de información, comercio, comunicación y conocimiento colectivo, y sus recomendaciones se consideran una norma a seguir. OWL es el lenguaje de ontologías que recomienda la W3C. Si acudimos al trabajo de McGuinness y van Harmelen (McGuinness y van Harmelen, 2004), el lenguaje de ontologías web (*Ontology Web Language, OWL*) está diseñado para ser usado en aplicaciones que necesitan procesar el contenido de la información en lugar de únicamente representar información para los humanos. OWL facilita un mejor mecanismo de interpretación de contenido Web que los mecanismos admitidos por XML, RDF, y RDF esquema (RDF-S) proporcionando vocabulario adicional junto con una semántica formal.

Existe una redefinición y extensión del lenguaje OWL que se denomina OWL 2 y es actualmente el estándar recomendado por la W3C (Hitzler et al., 2009). Las diferencias entre OWL y OWL 2 se deben a mejoras en la funcionalidad del lenguaje y en la capacidad de los elementos que se usan para representar el conocimiento. El lenguaje de representación de conocimiento OWL 2 mantiene la misma finalidad: está diseñado para formular, intercambiar y permitir el razonamiento con conocimiento sobre un dominio de interés.

OWL 2 utiliza una sintaxis que se denomina *RDF/XML* y está basada principalmente en el lenguaje RDF, que es otro lenguaje informático podríamos decir más general (Brickley y Guha, 2004). Mediante RDF/XML se establecen los elementos básicos para la descripción de un vocabulario o en este caso de una ontología. Podemos entenderlo como un “préstamo” entre lenguajes, el lenguaje OWL 2 utiliza algunas “palabras” prestadas del lenguaje RDF/XML para empezar a crear “frases” en su propio lenguaje. Además de RDF/XML, podemos utilizar otras sintaxis para crear nuestras ontologías, como la Manchester OWL 2 Syntax de la que hablaremos más adelante y que utilizamos en algunas figuras y ejemplos de este capítulo (§8.2).

En Ingeniería del Conocimiento, los lenguajes como OWL, no se consideran lenguajes de programación sino lenguajes declarativos. Un lenguaje declarativo es aquel que nos permite describir un estado de hechos, pero no lleva a cabo ningún razonamiento o cálculo (Hitzler et al., 2009). Por eso, para ‘razonar’ o para que un sistema analice la información que hemos representado en OWL tenemos que hacer uso de *razonadores*. Los razonadores nos permiten inferir información añadida sobre la descripción realizada mediante el lenguaje OWL.

OWL tampoco es un lenguaje prescriptivo como lo puede ser XML con respecto a cómo debe construirse sintácticamente un documento. De hecho, no permite forzar que un tipo de dato deba estar presente sintácticamente, lo que debe tenerse en cuenta cuando utilicemos este lenguaje. Tampoco podemos considerar OWL como un esquema de bases de datos, es verdad que un documento OWL almacena información y que existe cierta analogía entre la información declarada y el contenido de un esquema de base de datos; sin embargo también existen grandes diferencias entre ambos que a nivel técnico se basan en la semántica utilizada. Por ejemplo, si un dato no está presente en una base de datos, se considera normalmente falso, basándose en lo que se denomina *closed-world assumption* o suposición del mundo cerrado; en cambio, en el caso de un documento en OWL se considera ausente (pero ‘posiblemente’ cierto) basándose en la *open-world assumption* o suposición del mundo abierto (Drummond y Shearer, 2006).

7.1 Componentes del lenguaje OWL 2

Antes de profundizar en los elementos que componen OWL, debemos saber que se basa en tres nociones básicas (Motik et al., 2009) que se conocen con el nombre de *entidades*, *axiomas* y *expresiones*.

Las **entidades**¹⁰ en OWL 2 son los elementos del lenguaje que se utilizan para referirse a los ‘conceptos’, tal y como los concebimos desde la Terminología, que se van a representar del mundo real. Constituyen un componente necesario para construir la ontología ya que ésta se articula alrededor de éstos. En este caso, las entidades agrupan a todo el conjunto de conceptos que se puede utilizar para hacer referencia al mundo real y se expresa a través de elementos como *clases*, *individuals*, *datatype properties* y *object properties*¹¹. Estos elementos nos sirven para representar en lenguaje OWL 2 cada uno de los objetos, actividades, características o relaciones que deben mencionarse en la ontología.

Las **expresiones** son combinaciones de entidades (*classes*, *properties* o *individuals*) para formar descripciones más complejas a partir de descripciones básicas. Estas expresiones se realizan a partir de un conjunto de primitivas como son: conectores booleanos de tipo Y, O y NO, formas restringidas de cuantificación universal y existencial y auto-restricciones. A través de las expresiones construimos frases en lenguaje OWL 2 que combinan entidades y conectores para describir a otras entidades presentes en la ontología. Por ejemplo, para expresar que “una baldosa esmaltada tiene resistencia a la abrasión superficial con clase III, utilizaremos una expresión que incluirá la clase descrita, una propiedad que representa la característica, un valor para representar el valor ‘clase III’ y una forma restringida de cuantificación universal.

Los **axiomas** son declaraciones básicas que se expresan en una ontología OWL 2. Cuando construimos una ontología en OWL 2 podemos establecer declaraciones que se dan por verdaderas en todo el ámbito de la ontología y que puede afectar a cualquiera de los elementos que componen la ontología. En otras palabras, son verdades que se explicitan directamente en la ontología porque van a determinar cómo se debe entender el conjunto de la ontología. Por ejemplo, podemos declarar un axioma que indique que todos los seres humanos son mamíferos.

7.1.1 Entidades OWL 2

Dentro de las entidades OWL 2 encontramos tres grandes elementos que se denominan *individual*, *class* y *property* (Motik et al., 2009).

El elemento *class*¹² se utiliza en OWL 2 para definir un conjunto de instancias o clases de instancias que comparten las mismas propiedades a través de una misma categoría o etiqueta. Normalmente se utiliza para representar los conceptos que forman parte de la ontología (Hitzler et al., 2009). Por defecto, existe una clase denominada ‘Thing’ que actúa como superclase de

¹⁰ Cabe decir que el concepto de ‘entidad’ en este contexto de OWL2 no coincide con el concepto de ‘entidad’ al que estamos acostumbrados en Terminología: en OWL2 no se refiere a un tipo de concepto distinto a ‘característica’, ‘relación’ o ‘actividad’ sino que se entiende como el elemento que constituye el vocabulario de una ontología.

¹¹ En adelante y siempre que esto no produzca ambigüedad, nos referiremos al elemento *class* como clase, a los elementos *object property* y *datatype property* como propiedad, al elemento *individual* como instancia y al elemento *annotation* o *annotation property* como anotación

¹² En adelante, traduciremos el nombre del elemento *class* como ‘clase’, siempre que no se cree ambigüedad.

todas las clases. Además, hay una clase denominada 'Nothing' donde el razonador incluirá las clases definidas en la ontología que no estén bien expresadas o resulten inconsistentes. Por ejemplo, la clase 'humano' está situado bajo la clase 'Thing' (figura 21).

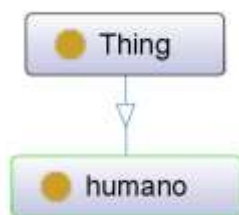


Figura 21. Clase 'humano' bajo la clase 'Thing'

Las clases se pueden organizar en una jerarquía de especialización mediante la expresión *subclassOf*. Por ejemplo, podemos decir "Humano *subclassOf* Mamífero" para expresar que un humano es un tipo de mamífero (figura 22).

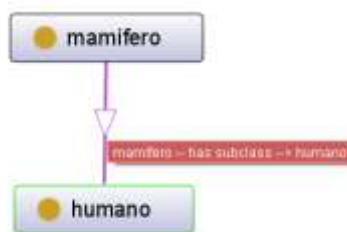


Figura 22. Clase 'humano' como subclase de 'mamífero'

El elemento *individual*¹³ hace referencia a una instancia de una clase, es decir, un miembro de una clase. Cabe remarcar que el elemento *individual* hace referencia a una instancia de cualquier clase de concepto y alude a un ejemplar concreto de un concepto (ibídem). Para establecer la relación entre una instancia y otra se utilizan elementos de tipo *property*. Si una instancia nueva compartiera las mismas *property* que una clase existente, la instancia se agruparía bajo esa clase si aplicáramos a la ontología un razonador. Por ejemplo, podemos decir que la clase 'humano' tiene un *individual* que es 'María' (figura 23).



Figura 23. Clase 'humano' y su instancia 'María'

El elemento *property*¹⁴ se utiliza para declarar una determinada propiedad o relación entre instancias o entre instancias y datos. El elemento *property* está diseñado para actuar en forma de

¹³ En adelante, traduciremos el nombre del elemento *individual* como 'instancia', siempre que no se cree ambigüedad.

¹⁴ En adelante, traduciremos el nombre del elemento *property* como 'propiedad', siempre que no se cree ambigüedad.

tripleto, y tendrá un dominio, una etiqueta y un rango. El dominio es el concepto (*class* o *individual*) que es descrito por la propiedad y el rango hace referencia al valor que se asigna al concepto descrito por la propiedad. Por ejemplo, en el tripleto ‘María’ ‘tieneHijo’ ‘Pau’, ‘María’ sería el dominio de la propiedad, ‘tieneHijo’ la etiqueta que da nombre a la propiedad y ‘Pau’ el valor.

En OWL 2, se distingue entre tres tipos de propiedades: *object property*, *datatype property* y *annotation property*.

El elemento *object property* se utiliza cuando la propiedad tiene como valor y rango una clase o instancia de la ontología. En cambio, cuando queremos representar un rasgo de una clase o instancia que tiene un valor que no está representado en la ontología a través de una clase o instancia se utiliza una *datatype property*. En este caso el valor del tripleto no será una clase o instancia de la ontología sino, por ejemplo, un número, un símbolo o un valor booleano. Los tipos de valores una *datatype property* se basan en las recomendaciones establecidas en XML Schema (Biron y Malhorta, 2004). Por ejemplo, *integer* se utiliza para indicar que el valor es un número entero, y será tratado como tal en los cálculos que se realicen.

Ejemplos de propiedades son ‘tieneHijo’, ‘tienePariente’, ‘tieneHermano’ o ‘tieneEdad’. Las tres primeras se puede utilizar para relacionar una instancia de una clase con otra y son ejemplos de una *object property* en OWL 2, por ejemplo: ‘María’ ‘tieneHijo’ ‘Juan’. En cambio, la propiedad ‘tieneEdad’ relaciona a una instancia de la clase ‘persona’ con un dato (su edad, un número entero) y se expresa mediante una propiedad de tipo *datatype property*, por ejemplo: ‘María’ ‘tieneEdad’ ‘28’. Para describir esto ‘María’ aparecerá en la ontología como *individual* pero ‘28’ no, porque no resulta rentable enumerar las posibles edades de una persona en forma de instancias. En su lugar, cuando describimos la *datatype property* indicamos solo el tipo de valor que puede adoptar, en este caso un número entero y ese valor concreto, 28, se expresará en la descripción de la instancia ‘María’ (figura 24).

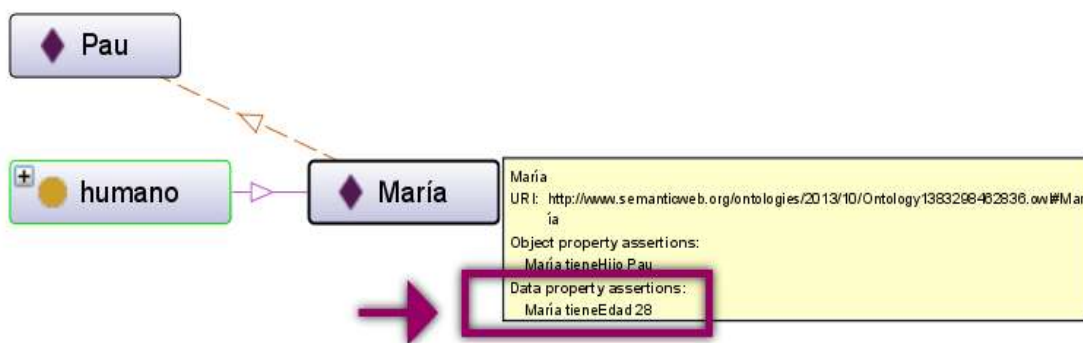


Figura 24. La instancia ‘María’ con la datatype property ‘tieneEdad’ y valor ‘28’.

Este modo de formalizar el valor numérico tiene sus ventajas de cara a aplicar razonadores en la ontología, ya que el valor podrá ser calculado por la máquina e interpretado.

En las *datatype property*, el valor del tripleto puede ser de varios tipos. El elemento *datatype* describe al conjunto de tipos de valor que puede tener como rango una *datatype property*. Normalmente, se utiliza para describir el tipo de valor que contiene o va asociado a una *datatype property*, entre ellos destacamos:

- *string*, indica que el tipo de dato es una cadena de caracteres, por ejemplo, una palabra que no representa a una clase o instancia en la ontología
- *boolean*, indica que el tipo de dato es booleano, un dato de valor binario, por ejemplo, verdadero o falso, sí o no.
- *double*, indica que el tipo de dato es un número en coma flotante IEEE 754 de 64 bits y precisión doble. Pueden ser positivos o negativos. Por ejemplo, -1.23 o 12678967.54.
- *integer*, indica que el tipo de dato es el conjunto de números enteros. Por ejemplo, -1, 0, 2, 1...
- *nonNegativeInteger* y *positiveInteger* refinan el tipo de número entero. En el primer caso incluyen los números enteros que no sean negativos incluido el cero y en el segundo caso los números positivos (1, 2, 3...).

Otros tipo de datos son *float*, *nonPositiveInteger*, *negativeInteger*, *long*, *int*, *short*, *byte*, *unsignedLong*, *unsignedInt*, *unsignedShort*, *unsignedByte*, *hexBinary*, *base64Binary*, *dateTime*, *time*, *date*, *YearMonth*, *gYear*, *gMonthDay*, *gDay*, *gMonth*, *anyURI*, *token*, *language*, *NMTOKEN*, *Name*, *NCName* (Biron y Malhorta, 2004).

Finalmente encontramos un tercer tipo de *property* en OWL 2 que se denomina ***annotation property***. Este tipo de propiedad sirve para añadir información sobre la ontología o los conceptos representados en la ontología a modo de anotación sin que se considere como parte de la estructura lógica de la ontología. Es decir, no aportan información lógica a la ontología. Normalmente, se usan para describir información contextual sobre una clase, instancia o propiedad. A través de una *annotation property* podemos incluir distintos tipos de información en la ontología, esta información se codifica a través de distintos campos que a su vez pueden incluir distintos tipos de dato (como los que hemos visto para las ***datatype property***).

Entre los campos existentes en OWL 2 encontramos ***rdfs:label***. Este campo puede servir para agregar otro nombre en lenguaje natural a un elemento presente en la ontología, este campo permite además vincular su contenido a una lengua en concreto. Por ejemplo, podemos agregar una anotación en la clase ‘humano’ mediante una *annotation property* cuyo campo *label* sea ‘human’@en (donde @en, expresa el idioma en el que está escrito el contenido del campo). De este modo, se nos permite indicar que la clase ‘humano’ tiene un nombre en inglés que es *human* (figura 25).



Figura 25. *Anotation property* en la clase ‘humano’ mediante *label* para indicar denominación equivalente en inglés.

Además, de este campo, a través de una *annotation property* podemos agregar información de otros tipos, como comentarios (rdfs:comment), referencias a otros conceptos (rdfs:seeAlso), etc.

En la estructura de las *property* en una ontología, existe la posibilidad de establecer jerarquías de forma similar a una estructura de clases. En el lenguaje OWL 2 se representa mediante *subPropertyOf* y se puede utilizar para inferir información mediante un razonador basándose en la jerarquía de *property*. Por ejemplo, 'tieneHijo' puede declararse como subpropiedad de 'tienePariente', por lo que se infiere que si María 'tieneHijo' a un miembro de la clase 'humano', 'María' también 'tienePariente' a un miembro de la misma clase (figura 26).

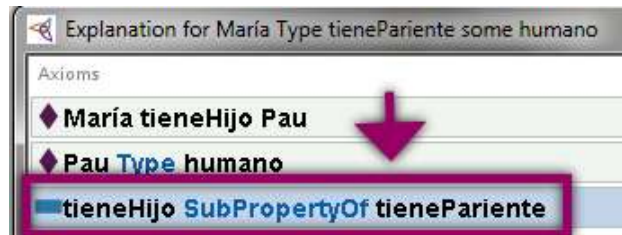


Figura 26. Inferencia de la clase 'María' en base a que 'tieneHijo' es subpropiedad de 'tienePariente'

7.1.2 Descripción de tipos de clases e instancias OWL 2

OWL 2 permite establecer jerarquías entre los conceptos que se representan en la ontología cuando se representan como *class* o como *property*. Y, a su vez, permite anidar las instancias (*individual*) a clases de la ontología, atendiendo a su descripción. Además, podemos expresar otro tipo de información para describir las relaciones entre los elementos de la ontología. A continuación, detallamos algunas posibilidades.

En el caso de querer expresar que dos clases, instancias o propiedades hacen referencia a un mismo concepto, OWL 2 nos proporciona etiquetas como *equivalentClasses* o *equivalentProperty* para indicar que se trata de un mismo concepto. En Manchester OWL Syntax, como se verá en el ejemplo, se expresa como *EquivalentTo*. Cuando dos clases son equivalentes, éstas tendrán compartidos miembros de clase. Por ejemplo, si declaramos que 'humano' y 'persona' son clases equivalentes, los miembros de 'humano' serán miembros de 'persona' y viceversa. En el ejemplo (figura 27), se muestra la explicación de cómo Pau que pertenece a la clase "humano" es también miembro de la clase "persona" al ser estas dos clases equivalentes:



Figura 27. Inferencia en la instancia 'Pau' en base a que la clase 'humano' equivale a la clase 'persona'

Igualmente, si 'tieneFamiliar' y 'tienePariente' son equivalentes y la instancia 'Juan' está relacionada con la instancia 'Pedro' mediante la propiedad 'tienePariente', se infiere que 'Juan' también estará relacionado con 'Pedro' mediante la propiedad 'tieneFamiliar':

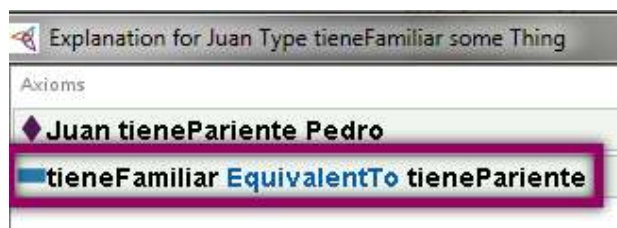


Figura 28. Inferencia en la instancia 'Juan' en base a que 'tieneFamiliar' es equivalente a 'tienePariente'

En el caso de *individuals*, la equivalencia también indica la existencia de distintos nombres para una misma instancia a través de *SameIndividual*. Por ejemplo, si 'María Soler' y 'María' son instancias equivalentes, se considera que son una misma instancia; las relaciones que se establezcan con la instancia 'María Soler', se establecerán también con la instancia 'María'.

En OWL 2, también podemos declarar que existe incompatibilidad entre elementos de la ontología. En este caso, podemos expresar distintos grados de incompatibilidad. En el caso de clases de la ontología, podemos indicar que una clase no se puede componer nunca de elementos que estén contenidos en otra clase y lo hacemos a través de *disjointWith*. Por ejemplo, podemos declarar que la clase 'hombre' y la clase 'mujer' son disjuntas de modo que una instancia de 'hombre' no pueda serlo de 'mujer'. En la figura 29, a continuación, se muestra la explicación del razonador de por qué Pau no es una mujer en base a que éste pertenece a la clase 'hombre', disjunta de la clase 'mujer'.

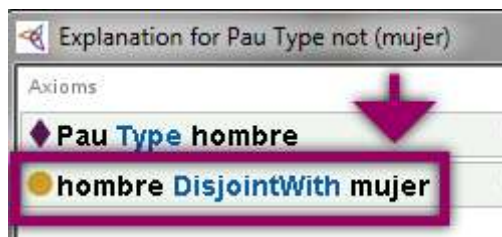


Figura 29. Explicación de por qué la instancia 'Pau' no pertenece a la clase 'mujer' en base a que es disjunta a 'hombre'

En el caso de las instancias, utilizamos *differentFrom* para explicitar que una instancia es distinta de otra. Otra alternativa que proporciona OWL 2 para declarar la diferencia entre instancias es *allDifferent*. Cuando queremos explicitar que una serie de instancias son distintas unas de otras, utilizamos *allDifferent*. Se diferencia del constructo *differentFrom* en que en este caso todos los miembros son distintos entre ellos. Por ejemplo, podemos explicitar que 'Juan' es *allDifferent* de 'María' y 'Elena', y estaríamos describiendo que 'María' y 'Elena' son también distintas entre ellas.

Es importante señalar que para describir las clases de una ontología en OWL 2, se suele recurrir a la combinación de elementos. Normalmente, se hace uso de dos opciones que combinan el elemento *equivalentClasses* o el elemento *subclassOf*. Estos elementos se completan con expresiones más complejas que combinan el uso de *property* para describir una clase. Por ejemplo, podemos describir que la clase 'adolescente' equivale a las instancias de la clase 'humano' que 'tieneEdad' entre 11 y 20 años:

```
Class: Adolescente
```

```
EquivalentTo: humano and (tieneEdad some integer [ >=11, <= 20])
```

Este tipo de expresiones son las que van construyendo la ontología en OWL 2 y sirven para describir las relaciones entre las clases, instancias y propiedades de la ontología.

7.1.3 Descripción de tipos de propiedades OWL 2

Cuando utilizamos una propiedad para describir una característica o relación entre dos conceptos de la ontología, la mayor parte del contenido descriptivo se realiza a través de la etiqueta en lenguaje natural que denomina a la propiedad. En OWL 2, existen mecanismos para delimitar la naturaleza o comportamiento de la propiedad de modo que podamos mejorar la capacidad descriptiva de ésta en la ontología. Por ejemplo, pongamos que la propiedad 'tieneHijo' vincula a la instancia 'María' con la instancia 'Juan'. La intención es decir que María es madre de Juan. En este punto, podríamos querer especificar que si 'María' 'tieneHijo' 'Juan', entonces 'Juan' no debería poder tener la propiedad 'tieneHijo' con la instancia 'María', porque un hijo no puede ser padre de su propia madre. En OWL 2 esto se realiza matizando la descripción de la propiedad. Para matizarla se recurre a expresar sus propiedades matemáticas, como las que hemos estudiado sobre las relaciones conceptuales en Terminología (§2.3.1). A continuación, describimos algunas posibilidades para delimitar la naturaleza de una propiedad:

- Relación inversa entre dos propiedades

Cuando necesitamos declarar que una propiedad es la inversa de otra, podemos hacer uso del elemento *inverseOf*. Si una propiedad P1 se ha declarado inversa de otra P2, cuando una clase o instancia X tenga la propiedad P1 con la clase o instancia V, V tendrá la propiedad P2 con X. En el ejemplo de María y Pau se muestra un ejemplo de inferencia del razonador a través de propiedades inversas. En la figura 30 se muestra cómo se ha inferido que Pau tiene la propiedad 'tienePadres' con valor 'María' en base a que la propiedad 'tieneHijo' es inversa de la propiedad 'tienePadres' y que 'María' tiene la propiedad 'tieneHijo' con la instancia 'Pau'.

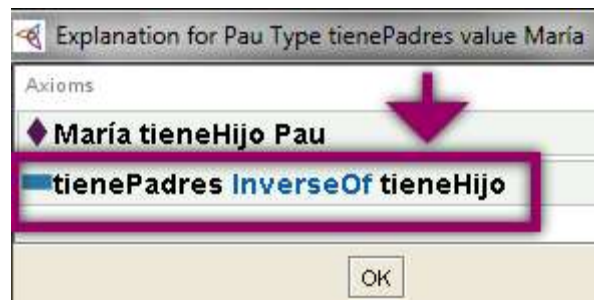


Figura 30. Inferencia de que la instancia 'Pau' tiene la propiedad 'tienePadres' con valor María en base a su inversa

Este elemento solo está disponible cuando describimos una característica o relación en forma de *object property*.

- Transitividad

Expresar que una propiedad es transitiva es algo necesario cuando queremos establecer vínculos de manera encadenada. Por ejemplo, si indicamos que la propiedad 'tieneAntepasado' es transitiva (mediante el elemento *transitiveProperty*) y vincula a 'María' con 'Pedro' y a 'Juan' con 'María', al haber expresado la transitividad un razonador de ontologías inferirá que

'Juan' 'tieneAntepasado' 'Pedro' (figura 31). Cuando una propiedad es transitiva, las clases o instancias vinculadas mediante ella de forma encadenada tienen la propiedad entre todos los miembros de la cadena. Este elemento solo está disponible cuando describimos una característica o relación en forma de *object property*.

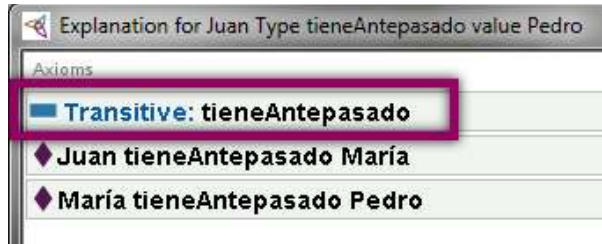


Figura 31. Inferencia entre instancias en base a una propiedad transitiva

- Simetría

La simetría entre propiedades tiene lugar cuando el rango y valor de una propiedad se puede invertir sin que la propiedad deje de tener sentido en esencia. En OWL 2, este rasgo se declara a través del elemento *symmetricProperty*. Este elemento solo está disponible cuando describimos una característica o relación en forma de *object property*.



Figura 32. Inferencia entre instancias en base a una propiedad simétrica

Por ejemplo, en la figura 32 se muestra la explicación del razonador de por qué se infiere que 'Javier' tiene la propiedad 'tieneAmigo' con valor 'Juan'. Dado que 'tieneAmigo' es una propiedad simétrica, si 'Juan' 'tieneAmigo' 'Javier', 'Javier' 'tieneAmigo' 'Juan'.

- Limitación de valor único

En el ejemplo de la introducción hemos hablado de la necesidad de declarar si una propiedad puede tener más de un valor cuando describe a una misma clase o instancia. A través del elemento *functionalProperty*, en OWL 2 podemos indicar que una propiedad solo puede tener un único rango cuando describe un mismo dominio. Por ejemplo, si la propiedad 'tieneID' es funcional y la instancia 'Juan' tiene la propiedad 'tieneID' con más de un valor, el razonador indicará que existe un error de consistencia en la ontología. En la figura 33, vemos el ejemplo.

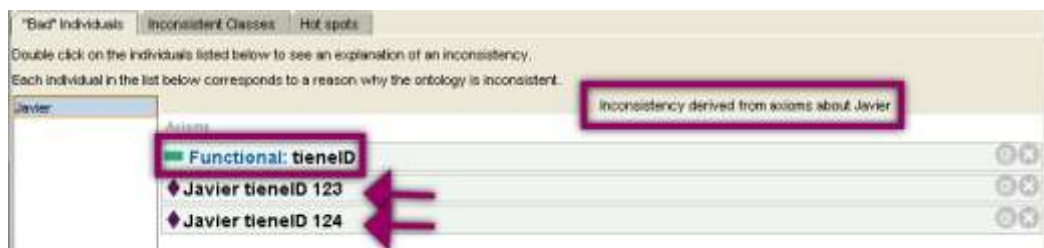


Figura 33. Inconsistencia detectada en la instancia en base a una propiedad funcional

En este caso, tanto una *object property* como una *datatype property* pueden definirse como funcionales.

En esta línea, podemos ir un paso más allá y definir una propiedad como *inverseFunctionalProperty*. Cuando una propiedad es inversa funcional, estamos señalando que es funcional y que su inversa también lo es. Por ejemplo, si 'tieneDNI' es inversa funcional de 'DNIde', entonces la instancia 'Juan que tiene la propiedad 'tieneDNI' solo podrá tener un rango y no puede haber ninguna otra persona que tenga ese mismo rango ya que 'DNIde' también es funcional. Solo las propiedades de tipo *object property* pueden describirse como *inverseFunctionalProperty*.

Hasta ahora hemos visto algunas de las posibilidades de OWL 2 para restringir y describir que rangos o valores puede tener una propiedad. A esto, podemos unir la posibilidad de restringir el número de veces que aparece una propiedad en la descripción de una misma clase. Podemos restringir el mínimo de veces posible que puede aparecer una propiedad en la descripción de un concepto de la ontología mediante *minCardinality*, en Manchester OWL Syntax *min*. Por ejemplo, la clase 'persona_amigable' se ha descrito como una clase que equivale a tener como mínimo una vez la propiedad 'tieneAmigo'. Por eso, tal y como se visualiza en la figura 34, la instancia 'Juan' se consideran como instancia de 'persona_amigable', ya que la descripción de esta instancia incluye la propiedad 'tieneAmigo' con valor 'Javier'.



Figura 34. Inferencia en base un axioma de equivalencia basado en una restricción de cardinalidad mínima

En el siguiente ejemplo se muestra la restricción de cardinalidad mínima de la clase "Personal amigable" Con *minCardinality*, en Manchester OWL Syntax *min*, si la propiedad 'tieneHijo' con respecto a la clase 'madre' tiene la restricción de cardinalidad mínima 1, cualquier instancia de la clase 'madre' tendrá como mínimo 1 relación con alguna instancia mediante la propiedad 'tieneHijo'. Si indicamos que una propiedad tiene cardinalidad mínima cero entonces esa propiedad es opcional.

El elemento que declara lo opuesto es *maxCardinality*. La cardinalidad máxima sirve para indicar que una propiedad con respecto a una clase ocurre un máximo determinado de veces. Si en la cardinalidad máxima indicamos cero, estamos diciendo que esa clase no puede tener esa propiedad nunca, por ejemplo, la clase 'soltero' tiene la propiedad 'tieneEsposa' con cardinalidad máxima cero (Hitzler et al., 2009). Finalmente, también podemos indicar ambos aspectos con el elemento *cardinality*. La restricción de cardinalidad se aplica a una propiedad cuando queremos indicar una misma cifra como su cardinalidad máxima y mínima. Si indicamos el mismo número en máxima y en mínima estaremos restringiendo el número exacto de veces que una propiedad puede tener en una relación con una determinada clase. Por ejemplo, si la propiedad 'tieneNombre' es mínimo 1 y máximo 1 para la clase 'persona' indicamos que una instancia de persona solo puede tener un nombre.

7.1.4 Descripción de conceptos a través de clases

Hasta el momento hemos visto las posibilidades de OWL 2 para declarar que algo es cierto o es falso. En este apartado, presentamos las opciones para establecer restricciones sobre el contenido de una clase, por ejemplo, qué condiciones se deben cumplir para que una instancia pueda ser miembro de una clase. De esta forma, definimos el grado de importancia de una u otra propiedad en la descripción de una clase. Por ejemplo, podemos definir si para pertenecer a una clase no basta con tener una determinada propiedad sino que además es necesario que el valor de la propiedad sea uno en concreto. El lenguaje OWL 2 permite utilizar tres restricciones que se declaran a través de los elementos *allValuesFrom*, *someValuesFrom* o *hasValue*, dependiendo de lo que queramos expresar.

Cuando describimos una clase a través de una propiedad, podemos querer indicar que esa propiedad tiene que tener un valor perteneciente a un rango concreto (a una clase concreta) y solo de ese tipo, para ello utilizamos *ObjectAllValuesFrom* o *DataAllValuesFrom* (según el tipo de propiedad que se esté usando) o expresado en Manchester OWL 2 Syntax a través *only*. Por ejemplo, si expresamos que la clase “hombre” agrupa a las clases e instancias con la propiedad “tieneGénero” “masculino” y solo “masculino” y al mismo tiempo, declaramos que la clase “varón” “tieneGénero” “masculino”, el razonador inferirá que la clase “varón” pertenece a la clase “hombre”, tal y como se muestra en la siguiente figura 35:

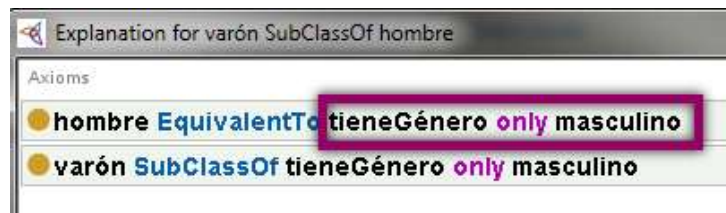


Figura 35. Inferencia de la relación entre las clases ‘hombre’ y ‘varón’ basada en la restricción del rango en la propiedad ‘tieneGénero’ que las describe

Esta es la forma en OWL 2 para delimitar el tipo valor que puede tener una propiedad cuando describe una clase.

Otra opción para restringir las propiedades cuando describen clases es *ObjectSomeValuesFrom* o *DataSomeValuesFrom* (según el tipo de propiedad que se esté usando) o expresado en Manchester OWL 2 Syntax a través *some*. Esta restricción expresa que una propiedad con respecto a una clase tiene como mínimo el rango declarado para esa propiedad (perteneciente a una determinada clase), aunque puede tener más. Esto implica que si esa propiedad está presente en una subclase o instancia de la clase, el rango de la propiedad debe ser el indicado al menos una vez. Por ejemplo, si en la clase ‘arroses’ declaramos tiene la propiedad ‘tieneIngrediente’ y como rango la clase ‘arroz’ con la restricción *ObjectSomeValuesFrom*, entonces si la clase ‘paella’ tiene como ingredientes ‘arroz’ se relacionará con la clase ‘arroses’ (figura 36).

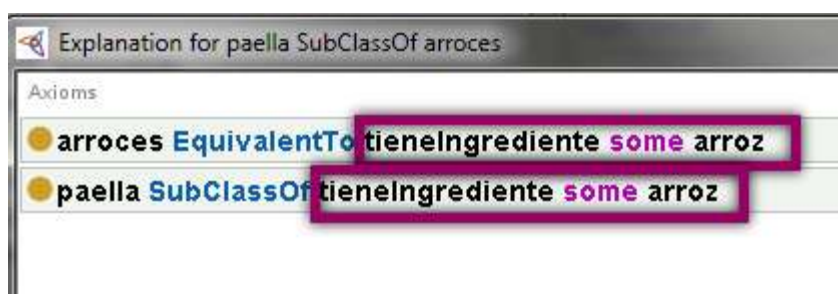


Figura 36. Inferencia de la relación entre las clases 'arroces' y 'paella' en base a la definición del rango de 'tieneIngrediente' que las describe

Finalmente, si lo queremos es indicar el valor concreto (representando en forma de instancia o dato) que toma una propiedad cuando describe una clase, OWL 2 proporciona la expresión *ObjectHasValue* o *DataHasValue* (según el tipo de propiedad que se esté usando), expresado en Manchester OWL 2 Syntax a través *value*. Este elemento servirá para restringir el valor de la propiedad de una clase a uno en concreto, que debe ser una instancia de una clase, y se denomina *valor de propiedad*. Por ejemplo, para expresar que los habitantes de Menorca son menorquines, creamos una propiedad llamada 'gentilicio' y en la descripción de la clase 'Menorca' indicamos que la propiedad 'tieneGentilicio' tiene como valor concreto una instancia llamada 'menorquín' que pertenece a la clase 'gentilicios'. En la figura 37, mostramos el resultado de la consulta con el razonador sobre los lugares con gentilicio 'menorquín' y el resultado de 'Menorca'.



Figura 37. Inferencia de la relación entre las clases 'Menorca' y 'lugares' en base a la definición del rango de 'tieneGentilicio' que las describe.

Finalmente, debemos mencionar que las instancias (*individual*) también pueden describirse directamente y no solo a partir de su(s) clase(s). En OWL 2, podemos describir instancias como dominio de una propiedad pero tenemos que tener en cuenta que los valores solo podrán ser otras instancias. Por ejemplo, podemos decir que 'Pedro' 'tieneAmigo' 'Juan', siendo 'Juan' y 'Pedro' instancias de la clase 'hombre', pero no podemos decir que 'Pedro' 'tieneAmigo' 'hombre'.

La descripción de una ontología en OWL 2 se basa en la combinación de elementos del lenguaje como los que hemos presentado en estos apartados y el nivel de complejidad en las expresiones puede llegar a abrumar al lector. Escribir directamente en OWL 2 requiere invertir un tiempo y un esfuerzo notables, y más si tenemos en cuenta que debemos tener unos conocimientos básicos de otros lenguajes como RDF que se utilizan en combinación con OWL 2. Afortunadamente, existen herramientas y sintaxis que simplifican enormemente esta tarea, a continuación veremos un ejemplo de sintaxis y en §II.10, dos herramientas que, mediante una interfaz gráfica intuitiva, nos ayudan en la creación de ontologías.

7.2 Manchester OWL 2 Syntax

Desde que OWL 2 se ha convertido en una recomendación de la W3C, este lenguaje se utiliza por parte de comunidades de expertos que no siempre son expertos en lógica ni tienen grandes conocimientos sobre su sintaxis o los axiomas a utilizar. Tal y como se desprende del apartado anterior OWL 2 puede resultar un lenguaje complejo y con una sintaxis enrevesada y poco escueta que dificulta su escritura. La sintaxis Manchester OWL 2 Syntax (Horridge y Patel-Schneider, 2009), desarrollada en la Universidad de Manchester, pretende evitar los problemas de complejidad que supone el lenguaje para un desarrollador no experto siendo fieles al lenguaje en sí. Se trata de una sintaxis que resulta más fácil de escribir y que puede utilizarse en herramientas de edición de ontologías de un modo sencillo (Horridge et al., 2006: 3).

Para llevar a cabo esta sintaxis, se centraron en la concisión y el uso de una simbología fácil de asimilar y utilizar. Primero, realizaron una serie de pruebas con usuarios de la herramienta Protégé-OWL, cuando todavía era un plug-in de Protégé-Frames (ibídem: 2), para ver cómo se desenvolvían los usuarios en la creación de ontologías con el lenguaje OWL. Horridge y otros se centraron en identificar cuáles eran las expresiones más utilizadas en la edición de clases para, por ejemplo, minimizar el uso de paréntesis necesarios en las expresiones o escoger palabras clave intuitivas que resultaran más adecuadas para el usuario final. El objetivo era crear una sintaxis que fuera más legible y ahorrara tiempo y esfuerzo sin despegarse de lo establecido en las especificaciones de lenguaje.

Existen otras sintaxis con la misma finalidad (Patel-Schneider et al., 2004) pero la ventaja de esta sintaxis con respecto a otras es que aprovecha las mejoras de otras sintaxis y las incorpora. En la Manchester OWL 2 Syntax, los símbolos matemáticos que se utilizaban para crear descripciones como \exists , \forall o \neg en algunas sintaxis basadas en DL han sido sustituidos por palabras como *some*, *only* o *not*.

Otra mejora ha sido utilizar la anotación por infijos en lugar de por prefijos ya que se observó que los usuarios no expertos en lógica a menudo no interpretaban correctamente la anotación por infijos. Por ejemplo, si en la descripción de la clase *pizza* se indicaba “ObjectSomeValuesFrom hasTopping MozzarellaTopping”, los no-expertos en lógica entendían “algunas pizzas tienen condimentos que son mozzarella” cuando en realidad debían entender “todas las pizzas tienen condimentos y algunos son mozzarella”.

A continuación, se muestran algunos ejemplos de elementos utilizados en la creación de ontologías y los cambios que supone utilizar la Manchester OWL 2 Syntax con respecto a OWL 2. Esta sintaxis está disponible en herramientas como Protégé-OWL 2 4.1.

7.2.1 Restricciones

Para hacer restricciones se utilizarán infijos en lugar de prefijos y los símbolos de la lógica de descripciones se reemplazarán por palabras en inglés. En la siguiente tabla podemos observar la construcción que se utiliza normalmente en OWL 2, el símbolo basado en la lógica de descripciones que se utilizaba en el programa Protégé-OWL 2, con el que hicieron las pruebas de usuario Horridge et al. (Horridge et al., 2006: ibídem), y la palabra que propone la Manchester OWL 2 Syntax como sustituto de estos tipos de representación, así como un ejemplo.

Construcción OWL 2	Símbolo en DL	Palabra en Manchester OWL 2 Syntax	Ejemplo
someValuesFrom	\exists	some	hasChild some Man
allValuesFrom	\forall	only	hasSibling only Woman
hasValue	\ni	value	hasCountryOfOrigin value England
minCardinality	\geq	min	hasChild min 3
cardinality	$=$	exactly	hasChild exactly 3
maxCardinality	\leq	max	hasChild max 3

Tabla 5. Construcciones de restricción en OWL2, en DL y en Manchester OWL 2 Syntax

7.2.2 Construcciones de clases booleanas

Las construcciones para definir clases mediante elementos booleanos también se han reemplazado por palabras equivalentes en inglés.

Construcción en OWL 2	Símbolo en DL	Palabra en Manchester OWL 2 Syntax	ejemplo
intersectionOf	\sqcap	and	Doctor and Female
unionOf	\sqcup	or	Man or Woman
complementOf	\neg	not	not Child

Tabla 6. Construcciones booleanas en OWL 2, DL y Manchester OWL 2 Syntax

7.2.3 Expresiones de clases complejas

Las expresiones de clases complejas, o dicho de otro modo, expresiones que utilizan subordinaciones en la descripción, se pueden construir utilizando construcciones de clases booleanas. Por ejemplo: “Person and hasChild some (Person and (hasChild only Man) and (hasChild some Person))”, describe al conjunto de personas que tiene al menos un hijo que tiene hijos que son solo hombres, es decir, abuelos y abuelas que solo tienen nietos y no nietas. Los paréntesis se utilizan para desambiguar el significado de la expresión.

7.2.4 Valores de datos y propiedades *datatype*

Los valores de datos pueden ser un tipo, por ejemplo, booleanos, enteros o decimales, o directamente un valor concreto, por ejemplo, 21. Los tipos dependen de la herramienta pero normalmente se basan en las recomendaciones establecidas en XML Schema (Biron y Malhorta, 2004).

Las constantes se pueden expresar sin explicitar el tipo, simplemente entre comillas dobles, o indicando el tipo a continuación. Por ejemplo, “hasAge value "21"^^integer”, tiene 21 años, siendo 21 un tipo de valor entero. El uso de tipos de datos en expresiones más generales se puede realizar utilizando el nombre abreviado del tipo. Por ejemplo, “hasAge some int” para número entero, tiene edad: un número entero.

Además, podemos agregar descriptores adicionales para crear nuevos tipos de datos. Por ejemplo, “Person and hasAge some int[>= 65]” indica que el valor será un número entero mayor o igual a 65.

También podemos utilizar descriptores múltiples. Por ejemplo, para expresar rangos numéricos, “Person and hasAge some int [>= 18, <= 30]”, una persona que tiene entre 18 y 30 años.

7.3 Recapitulación y reflexión

En resumen, todos estos elementos aquí presentados constituyen la base de la descripción de una ontología, estos elementos se utilizarán para construir la estructura que compone toda la ontología. En este apartado hemos explorado las características del lenguaje OWL 2 , el lenguaje por excelencia para la creación de ontologías. Para un usuario no experto en lenguajes informáticos ni lógicos, este lenguaje resulta complejo, sin embargo, podemos abstraernos de los formalismos del lenguaje y centrarnos en los principales aspectos que representa y por los que estamos presentado el lenguaje OWL 2 en este trabajo.

OWL 2 es un lenguaje informático que sirve para crear ontologías, este lenguaje permite diferentes niveles de expresión pero conviene considerar que según aumentemos la expresión del lenguaje reducimos las posibilidades de que un ordenador pueda procesar la información expresada. Por eso, un buen criterio para basar la creación de una ontología es hacer uso de los elementos aptos para OWL 2 DL, que constituye el término medio entre la expresión y la representación formal procesable.

En OWL 2 existen tres grandes tipos de elementos:

1. los elementos que describen la naturaleza de una clase o instancia
2. los elementos que describen la naturaleza de una propiedad
3. los elementos que describen la relación entre una propiedad y una clase

Las clases hacen referencia a conceptos que sirven para agrupar a otros conceptos que comparten las características de la clase (subclases o instancias). Una instancia es un ejemplar de un concepto, el referente y no puede tener sub-instancias por debajo. Una propiedad OWL 2 es un elemento que permite establecer vínculos entre clases o instancias. Si el vínculo relaciona a un concepto (clase o instancia) con un tipo de dato se tratará de una *datatype property*, en cambio si la relación vincula a un concepto (clase o instancia) con otro u otros en la ontología se tratará de una *object property*.

Como vemos, el concepto de ‘característica’ y el de ‘propiedad OWL 2 ’ no son del todo equivalentes, ya que la propiedad para el sistema solo significa etiqueta que describe la relación entre dos elementos que son descritos. Se podría equiparar el rol de elemento descriptor de una característica al papel que cumple una propiedad OWL 2 pero este elemento no es suficientemente complejo como para describir también la faceta de concepto de una característica y menos aún todas las propiedades de ésta. No obstante, apreciamos que este elemento será muy útil para ayudar a describir una característica en la ontología y su papel como elemento descriptor.

Finalmente, nos gustaría cerrar este capítulo sobre OWL 2, haciendo hincapié en la utilidad y expresividad de este lenguaje para describir conceptos, como se puede intuir

observando los ejemplos de este apartado, y la complejidad a la que puede llegar este lenguaje en términos de representación. Además, valoramos muy positivamente la existencia de ‘lenguajes intermedios’ como la Manchester OWL 2 Syntax que favorecen y acercan el lenguaje OWL 2 a usuarios no expertos.

8 Metodologías de representación de ontologías: methontology

Methontology es una metodología para la construcción de ontologías a nivel de conocimiento desarrollada por el equipo *Ontology Engineering Group* que dirige Gómez-Pérez desde la Universidad Politécnica de Madrid¹⁵. *Methontology*, a diferencia de otras metodologías, se propone como una metodología general y se inspira en procesos de desarrollo de software y metodologías de la Ingeniería del Conocimiento (Gómez-Pérez y Fernández-López, 2004: 125).

En *Methontology* se identifican distintas actividades durante el proceso de desarrollo, estas actividades se agrupan en tres grandes grupos: actividades de gestión, actividades de apoyo y actividades de desarrollo (ibídem: 109). En la siguiente tabla, resumimos en qué consiste cada una de las actividades:

Gestión	Planificación: identificar las tareas a realizar y planificar su orden, el tiempo a invertir y los recursos necesarios.
	Seguimiento: comprobar que se está cumpliendo la planificación de forma adecuada.
	Control de calidad: asegurar la calidad de cada uno de los resultados.
Apoyo	Adquisición de conocimiento: obtener conocimiento de los expertos a través de algún sistema semi-automático.
	Integración y/o fusión, y alineación: consiste en integrar una ontología antigua a la nueva y/o fusionar dos o más ontologías ya existentes. En consecuencia, habrá que alinear mediante enlaces o mapeado las ontologías involucradas.
	Evaluación: evaluar técnicamente ontologías existentes en el dominio, el entorno de software y la documentación a utilizar.
	Documentación: detallar, de forma clara y exhaustiva, cada una de las fases completadas y los resultados obtenidos.
	Gestión de la configuración: guardar todas las versiones de la documentación y de la ontología para controlar los cambios.
Desarrollo	Especificación: declarar porqué se construye la ontología, para qué se va a utilizar, quién es el usuario final.
	Conceptualización: organizar y convertir la información de la que disponemos sobre el dominio en una especificación semi-formal que sirva de modelo conceptual de la ontología. El resultado de la conceptualización es el propio modelo conceptual de la ontología que se formalizará transfiriendo este modelo a un lenguaje formal.
	Formalización: convertir el modelo conceptual en una especificación formal computable mediante un lenguaje de ontologías.

¹⁵ <http://www.oeg-upm.net/>

	Implementación: construir todos los modelos computables necesarios en el lenguaje de ontologías
	Mantenimiento: actualizar y corregir la ontología si es necesario.

Tabla 7. Actividades durante el proceso de desarrollo de una ontología (Gómez-Pérez y Fernández-López, 2004: 109)

Methontology propone un ciclo de vida de la construcción de la ontología basado en prototipos que evolucionan (fig. 38). Esto permite añadir, modificar y eliminar términos en cada versión o prototipo.

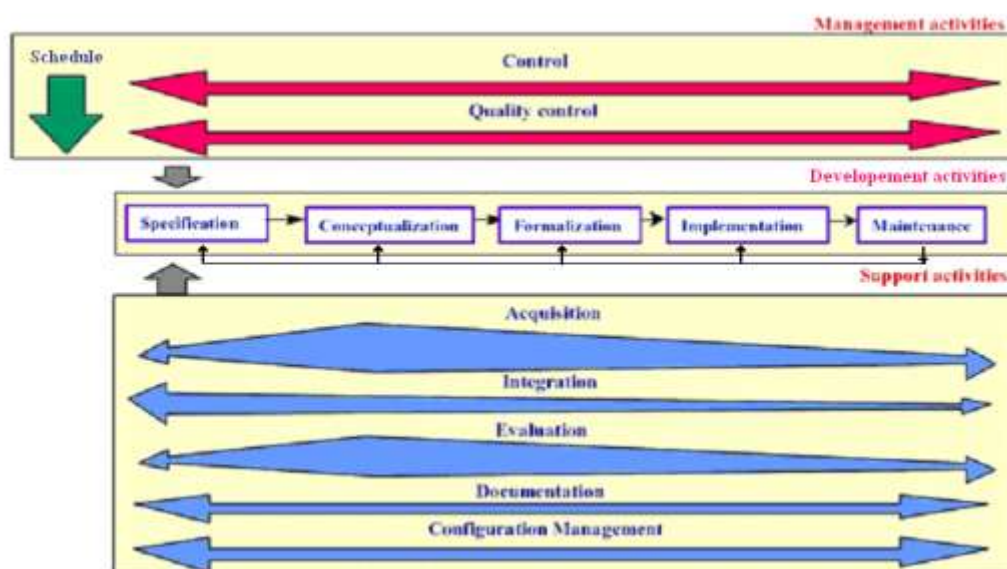


Figura 38. Proceso de desarrollo y ciclo de vida de *Methontology* (ibídem: 127)

El primer paso del ciclo se centra en la actividad de planificación. A continuación, empieza la actividad de especificación que arranca al mismo tiempo que las otras actividades de gestión (control de calidad y seguimiento), las de apoyo (adquisición de conocimiento, integración, evaluación, documentación y gestión de la configuración) y las actividades de desarrollo (especificación, conceptualización, formalización, implementación y mantenimiento). Es decir, los tres grupos de actividades se van llevando a cabo en paralelo, salvo en el primer grupo la actividad de planificación. Dentro de cada grupo las actividades de gestión y apoyo se realizan en paralelo pero las de desarrollo se realizan de forma consecutiva.

En lo referente a las actividades de desarrollo, una vez se ha especificado el primer prototipo, se elabora el modelo conceptual durante la tarea de conceptualización. La formalización y la implementación se suceden a continuación una detrás de otra. Si utilizamos una herramienta de edición de ontologías, el modelo de conceptualización se puede implementar directamente en distintos lenguajes de ontologías mediante *traductores*. Por esta razón, en *Methontology*, la actividad formalización no siempre es obligatoria (Gómez-Pérez, Fernández-López, 2004:126).

Finalmente, terminamos nuestro recorrido por *Methontology* centrándonos en la fase de conceptualización, donde se recomienda:

- Tarea 1: Construir el glosario de términos

Se trata de elaborar un listado de todos los términos y palabras que van a aparecer en la ontología. Esta tarea se realizará con la ayuda de los expertos, el ingeniero de conocimiento y los lingüistas. En esta fase, los desarrolladores de la ontología anotarán los términos que son sinónimos. Un ejemplo de los campos que debe contener el glosario es: nombre, sinónimos, acrónimos, descripción y tipo. *Methontology* distingue los tipos según su rol en la representación de conocimiento: conceptos, atributos, relaciones, instancias, etc.

- Tarea 2: Construir taxonomías de conceptos

Para construir taxonomías de conceptos, se seleccionan del glosario de términos aquellos que serán más relevantes para el dominio. *Methontology* propone utilizar las cuatro relaciones taxonómicas definidas en *Frame Ontology* y en OKBC (ibídem: 48): subclase-de, descomposición-disjunta, descomposición-exhaustiva y partición.

Un concepto C1 es subclase-de otro concepto C2 si y solo si, todas las instancias de C1 lo son también de C2. La descomposición-disjunta se utiliza para indicar que una clase se compone de subclases que no pueden tener una instancia en común entre ellas, aunque puede haber una instancia que no pertenezca a ninguna subclase. Por ejemplo, la clase *Vuelos British Airways* se compone de los vuelos *B0066*, *B0068* y *B0069*, pero si un vuelo es *B0066* ya no puede ser *B0068* ni *B0069*. La descomposición-exhaustiva señala que no puede haber una instancia de un concepto que no pertenezca al menos a una de sus subclases. Por ejemplo, los conceptos *económico*, *empresa* y *primera* son una descomposición exhaustiva de *paquete vacacional*, no puede haber ninguna instancia de *paquete vacacional* que no pertenezca al menos a una de esas clases, pudiendo pertenecer a más de una. La partición es similar a la descomposición-disjunta, solo que en ésta, una instancia debe pertenecer a una de las subclases obligatoriamente.

- Tarea 3: Construir diagramas de relaciones binarias ad hoc

Una vez construida, revisada y evaluada la taxonomía, el siguiente paso es construir diagramas de relaciones binarias ad hoc. El objetivo de este diagrama es señalar las relaciones ad hoc existentes entre conceptos de la misma o de distintas taxonomías de conceptos. Un ejemplo sería:

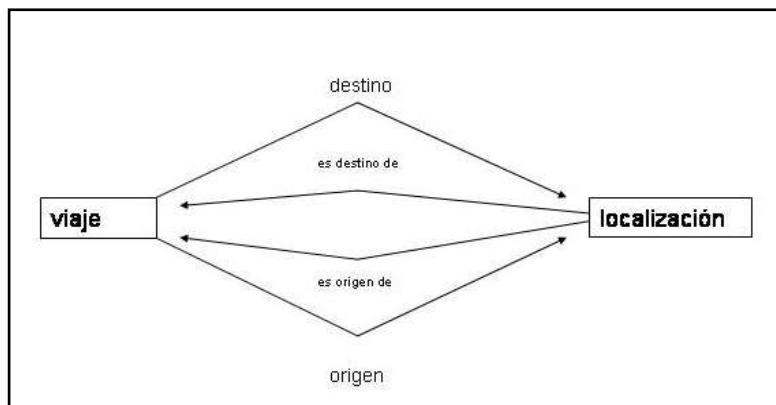


Figura 39. Ejemplo de diagrama de relaciones binarias

- Tarea 4: Construir el diccionario de conceptos

Una vez que las taxonomías de conceptos y los diagramas de relaciones binarias ad hoc han sido generados, el desarrollador de la ontología debe especificar cuáles son las propiedades que describen cada concepto de la taxonomía, así como las relaciones identificadas en los diagramas y las instancias de cada uno de los conceptos.

El diccionario de conceptos (tabla 8) contiene todos los conceptos del dominio, sus relaciones, sus instancias, y sus atributos de clase y de instancia. Las relaciones especificadas para cada concepto son aquéllas en las que el concepto es el origen de la misma. Por ejemplo, el concepto juicio es el origen de tres relaciones ad hoc: “demandante”, “demandado” y “se celebra en”. Las relaciones y los atributos de instancia y de clase son *locales* en los conceptos, lo que significa que sus nombres pueden aparecer repetidos en diversos conceptos.

Nombre del concepto	Atributos de clase	Atributos de instancia	Relaciones
AA7462	--	--	mismo vuelo que
vuelo American Airlines	nombre empresa	--	--
vuelo British Airways	nombre empresa	--	--
hotel 5 estrellas	número de estrellas	--	--
vuelo	--	--	mismo vuelo que
localización	--	nombre tamaño	es origen de es destino de
alojamiento	--	precio de la habitación estándar	situado en
viaje	--	fecha de llegada nombre de empresa fecha de salida tarifa ida y vuelta tarifa solo ida	destino origen
paquete vacacional	--	presupuesto precio final nombre número de días restricciones del viaje	destino salida alojado en viaja en
localización USA	--	--	--

Tabla 8. Ejemplo de diccionario de conceptos

- Tarea 5: Definir las relaciones binarias ad hoc

El objetivo de esta tarea es describir en detalle todas las relaciones binarias ad hoc identificadas en el diagrama de relaciones binarias e incluidas en el diccionario de conceptos.

Para cada relación binaria ad hoc, el desarrollador de la ontología debe especificar su nombre, los nombres de los conceptos origen y destino, cardinalidad (cuántas relaciones de ese tipo pueden establecerse para un mismo concepto), propiedades matemáticas, relación simétrica, transitiva, etc., y su relación inversa, si existe (tabla 9).

nombre relación	concepto origen	card. origen	concepto destino	propiedades matemáticas	relación inversa
mismo vuelo que	vuelo	N	vuelo	simétrica transitiva	--
situado en	alojamiento	1	localización	--	--
alojado en	paquete vacacional	N	alojamiento	--	--
Origen	viaje	1	localización	--	es origen de
Destino	viaje	1	localización	--	es destino de
Origen	paquete vacacional	1	localización	--	es origen de
Destino	paquete vacacional	1	localización	--	es destino de

Tabla 9. Ejemplo de tabla de relaciones binarias *ad hoc*

- Tarea 6: Definir los atributos de instancia en detalle

El objetivo de esta tarea es describir en detalle todos los atributos de instancia incluidos en el diccionario de conceptos (tabla 10). Cada fila de la tabla de atributos de instancia contiene la descripción detallada de un atributo de instancia. Como ya se comentó anteriormente, los atributos de instancia describen a las instancias del concepto y sus valores pueden ser distintos para cada una de dichas instancias. Por cada atributo de instancia, el desarrollador de la ontología debe especificar su nombre, el concepto al que el atributo pertenece (los atributos son locales en los conceptos), tipo de valor, unidad de medida, precisión y rango de valores (en el caso de atributos numéricos), y cardinalidad mínima y máxima, expresada mediante la siguiente forma (mínimo de relaciones, máximo de relaciones).

nombre atributo instancia	nombre concepto	tipo valor	unidad medida	precisión	valor de rango	cardinalidad
presupuesto	negocios	decimal	moneda	0.01	1000...3000	(0,1)
presupuesto	económico	decimal	moneda	0.01	1000...3000	(0,1)
nombre	localización	cadena	--	--	--	(1,N)

Tabla 10. Ejemplo de tabla de atributos de instancia

En el caso de atributos cuyo tipo de valor es numérico, se suele integrar alguna ontología de unidades de medida. *Methontology* propone integrar ontologías durante la actividad de conceptualización, y no posponer dicha integración para la actividad de implementación de la ontología.

- Tarea 7: Definir los atributos de clase en detalle

El objetivo de esta tarea es describir en detalle todos los atributos de clase incluidos en el diccionario de conceptos. Para cada atributo de clase, el desarrollador de la ontología debe

rellenar la siguiente información: nombre del atributo, nombre del concepto donde el atributo se define, tipo de valor, unidad de medida (si es un valor numérico), cardinalidad y valor(es). Por ejemplo, (tabla 11):

nombre atributo	definido en concepto	tipo valor	unidad de medida	precisión	cardinalidad	valores
nombre empresa	Vuelo American Airlines	cadena	--	--	(1,1)	AA
nombre empresa	Vuelo British Airways	cadena	--	--	(1,1)	BA
nombre empresa	Vuelo Iberia	cadena	--	--	(1,1)	IB
número de estrellas	hotel 5 estrellas	entero	estrella	1	(1,1)	5
número de estrellas	hotel 4 estrellas	entero	estrella	1	(1,1)	4
número de estrellas	hotel 3 estrellas	entero	estrella	1	(1,1)	3
número de estrellas	hotel 2 estrellas	entero	estrella	1	(1,1)	2
número de estrellas	hotel 1 estrella	entero	estrella	1	(1,1)	1

Tabla 11. Ejemplo de tabla de atributos de clase

- Tarea 8: Definir las constantes en detalle

El objetivo de esta tarea es describir en detalle cada una de las constantes identificadas en el glosario de términos (tabla 12). Para cada constante, el desarrollador de la ontología debe especificar su nombre, tipo de valor, valor y unidad de medida en el caso de constantes numéricas.

Nombre	tipo valor	valor	unidad de medida
número máximo pasajeros en avión	entero	200	persona

Tabla 12. Ejemplo de tabla de constantes

- Tarea 9: Definir axiomas¹⁶ formales

Para realizar esta tarea, el desarrollador de ontologías debe identificar los axiomas formales que son necesarios en la ontología y describirlos de manera precisa. Para cada definición de axioma formal, *Methontology* propone especificar la siguiente información: nombre, descripción en lenguaje natural, expresión lógica que define de manera formal el

¹⁶ Un axioma es un postulado que se considera verdadero sin necesidad de demostración. En este contexto, los axiomas se pueden entender como las verdades sobre las cuales se construye la ontología.

axioma usando lógica de primer orden, y los conceptos, atributos y relaciones ad hoc utilizadas en el axioma así como las variables utilizadas. Un ejemplo de tabla que describe un axioma formal en una ontología de viajes es:

nombre axioma	tren en Europa
Descripción	todo tren que sale de una localización en Europa llega a otra localización en Europa
Expresión	para todo (?X, ?Y, ?Z) [viaje tren] (?X) y [origen] (?X, ?Y) y [destino] (?X, ?Z) y [localización europea] (?Y) -> [localización europea] (?Z)
Conceptos	viaje tren localización europea
atributos mencionados	--
relaciones binarias <i>ad hoc</i>	origen destino
Variables	?X ?Y ?Z

Tabla 13. Ejemplo de axioma formal

- Tarea 10: Definir reglas¹⁷

De manera similar a la tarea previa, en esta tarea el desarrollador de la ontología debe identificar en primer lugar qué reglas se necesitan en la ontología, y entonces describirlas en la tabla de reglas. Para cada regla, *Methontology* propone incluir la siguiente información: nombre, descripción en lenguaje natural, expresión que describe formalmente la regla, y conceptos, atributos y relaciones ad hoc utilizados en la regla, así como las variables usadas.

Methontology propone especificar las expresiones de las reglas utilizando el formato *si <condiciones> entonces <consecuencia>*. La parte izquierda de la regla es una conjunción de condiciones simples, mientras que la parte derecha es una simple expresión de un valor de la ontología. Un ejemplo de tabla de una regla:

nombre regla	regla Costa Cruceros
Descripción	todo barco que sale de Europa se contrata a la empresa Costa Cruceros

¹⁷ Una regla es un esquema que establece que si se da una condición tiene que darse una consecuencia. En este contexto, servirán para inferir información nueva a partir de la información representada.

Expresión	si [localización europea] (?Y) y barco (?X) y [origen] (?X,?Y) entonces [nombre empresa] (?X, “Costa Cruceros”)
Conceptos	barco localización europea
atributos mencionados	nombre empresa
relaciones binarias <i>ad hoc</i>	origen
VARIABLES	?X ?Y

Tabla 14. Ejemplo de tabla de una regla

- Tarea 11: Describir instancias

Una vez que el modelo conceptual de la ontología ha sido creado, se pueden definir las instancias que aparecen en el diccionario de conceptos. Para cada instancia se define: su nombre, el nombre del concepto al que pertenece, los atributos y los valores de los atributos de instancia (tabla 15).

nombre de instancia	nombre concepto	atributos	valores
AA7462_Feb08_2002	AA7462	nombre empresa	American Airlines
		fecha salida	08/02/02
		fecha llegada	08/02/02
		tarifa solo ida	300
AA7462_Feb16_2002	AA7462	nombre empresa	American Airlines
		fecha salida	16/02/02
		fecha llegada	16/02/02
		tarifa solo ida	300

Tabla 15. Ejemplo de tabla de instancias

El equipo *Ontology Engineering Group* ha utilizado esta metodología para construir ontologías de dominio con herramientas como ODE (Blázquez et al., 1998) y WebODE (Arpírez et al., 2003). Además, existen otros proyectos que han utilizado esta metodología con Protégé-2000 (Noy et al., 2000) u OntoEdit (Sure et al., 2002).

8.1 Recapitulación y observaciones

Desde nuestro punto de vista, conocer la propuesta de *Methontology* supone una parada necesaria para todo desarrollador de una ontología. Como hemos visto, en esta metodología no solo se refleja el trabajo del programador informático sino también el papel que los terminólogos o lingüistas deben jugar en la creación de una ontología. De esta metodología, podemos extraer las ideas principales que puestas en común con otras ideas de metodologías

vistas anteriormente (§4.2) determinarán la metodología utilizada en nuestro trabajo de creación e implementación de la ontología. Desde nuestro punto de vista, una de las principales ventajas de *Methontology* es que permite no abordar un determinado aspecto como ellos sugieren, adaptando el resto del planteamiento de modo que resulte compatible con los objetivos del trabajo y continúe aportando ideas creativas a nuestra forma de hacer y trabajar. Desde nuestro punto de vista, esto se consigue al haber diferenciado las distintas tareas que pueden llevarse a cabo en la creación y el diseño de la ontología. De esta forma, la metodología para realizar una tarea no interfiere en la siguiente, siempre que el resultado sea acorde a lo esperado.

9 Herramientas para desarrollar ontologías

Como mencionábamos anteriormente, OWL se ha utilizado en numerosas ontologías y existen aplicaciones como Protégé (Horridge, 2011), SWOOP (Kalyanpur et al., 2006), WebODE (Arpírez, Corcho et al., 2003) que ofrecen una interfaz para desarrollar ontologías basadas en OWL. Desafortunadamente, la mayoría de herramientas se desarrollaron para trabajar con OWL y a día de hoy, las herramientas disponibles más conocidas para desarrollar ontologías en OWL 2 son Protégé 4.1 (Horridge, 2011) y TopBraid Composer (Topquadrant, 2011). A continuación, vamos a describir estas dos herramientas, centrándonos especialmente en Protégé 4 ya que es la herramienta que hemos utilizado en nuestro trabajo empírico. Concretamente, en nuestro trabajo hemos utilizado la versión de Protégé 4.1. Además, cerraremos este apartado presentando otro tipo de programas que suelen combinarse con los de edición para aprovechar al máximo la información representada, se trata de los razonadores, que nos ayudan en la tarea de inferir información.

9.1 Protégé 4

Protégé 4 es una nueva versión de la aplicación Protégé basada en OWL 2. Esta herramienta ha sido desarrollada en la Universidad de Stanford en el *Stanford Center for Biomedical Informatics Research*¹⁸ y está disponible de forma gratuita bajo la licencia de código abierto MPL.

El editor de ontologías Protégé 4 permite, entre otros:

- Cargar y guardar ontologías OWL 2 entre otros formatos.
- Editar y visualizar clases, propiedades y reglas basadas en SWRL (un lenguaje de reglas orientado a la web semántica).
- Definir características de clases lógicas en forma de expresiones OWL 2
- Ejecutar razonadores para, por ejemplo, hacer clasificaciones basadas en la lógica descriptiva
- Editar instancias OWL 2 para marcado.

Protégé 4 mantiene algunos de los principios que tenía su antecesora Protégé2000 (Noy y McGuinness, 2005), permitiendo la construcción de ontologías de dominio, e incorpora muchas novedades, principalmente motivadas por su capacidad para trabajar con ontologías utilizando el lenguaje OWL 2, algo de lo que carecían las versiones anteriores. Esta adaptación al lenguaje OWL 2 conlleva también, no obstante, algunas diferencias negativas de usabilidad con respecto a versiones anteriores de Protégé. Principalmente, destaca la ausencia de una pestaña para elaborar formularios de introducción de datos, aunque esta carencia en cierta medida puede suplirse con la instalación de un complemento de tipo *plug-in*.

Al mismo tiempo, entre las ventajas de esta herramienta con respecto a otras destaca su naturaleza libre y gratuita, el soporte técnico (parcialmente gratuito), la gran variedad de aplicaciones de tipo *plug-in* gratuitas que la comunidad de usuarios ha desarrollado para mejorar la aplicación y, fundamentalmente, la posibilidad de elaborar ontologías totalmente compatibles con el lenguaje OWL 2, así como el uso de varios razonadores específicos en la misma interfaz.

¹⁸ La página del centro se puede consultar en <http://protege.stanford.edu>

Protégé 4 es un programa complejo con numerosas funcionalidades y opciones (Horridge, 2011). A continuación vamos a explorar, brevemente, algunas de las funcionalidades más importantes centrándonos en los aspectos básicos que conlleva la tarea de creación de una ontología de dominio. Para facilitar el seguimiento de la descripción, vamos a utilizar como ejemplo la creación de una ontología de pizzas similar a la que figura en el tutorial de la aplicación (Drummond et al., 2007; Horridge, 2011).

9.1.1 Interfaz del programa

La interfaz de Protégé 4 se compone, principalmente, de una barra superior de menú, una barra de navegación (similar a la de un navegador de internet) y un conjunto de ventanas superpuestas en forma de pestaña que puede ampliarse o reducirse según las necesidades del usuario (figura 40).

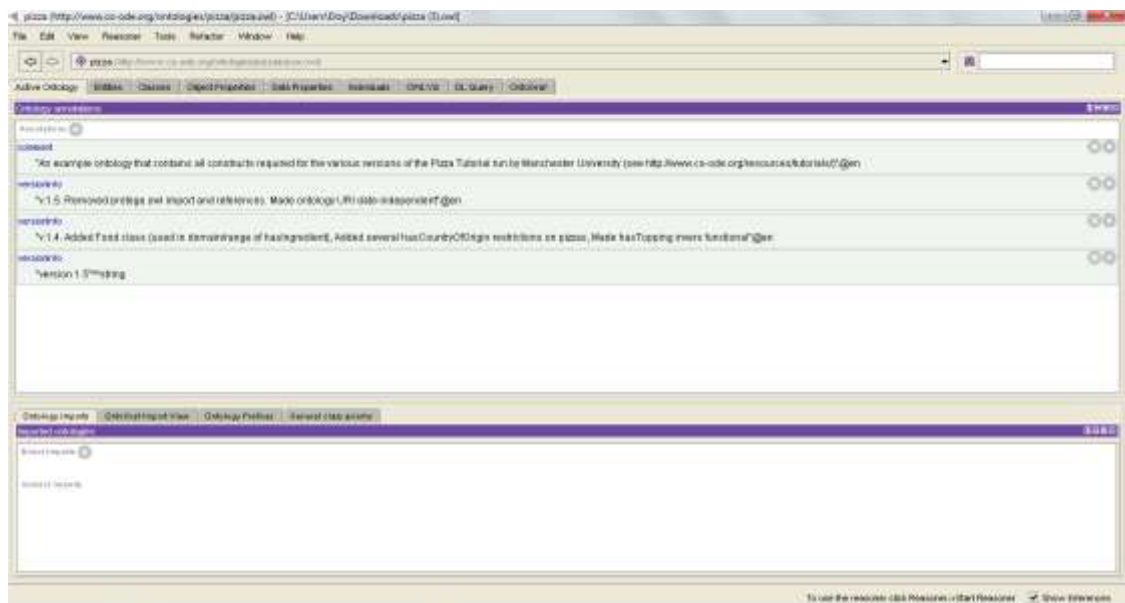


Figura 40. Interfaz de Protégé 4 con la ontología *pizza* abierta

La barra de menú, incluye el acceso a opciones generales de administración. Desde aquí, podemos realizar distintas tareas como crear o abrir una ontología, agregar clases de forma masiva, modificar las propiedades generales de una rama de la ontología, activar y desactivar el razonador, activar o desactivar complementos de tipo *plug-in*, modificar la visualización de la interfaz o agregar sub-ventanas con otros datos sobre la ontología (codificaciones en otros lenguajes, visualización de descripciones de entidades, etc.).

La barra de navegación nos muestra la ontología con la que estamos trabajando y nos permite hacer búsquedas de palabras que pueda contener la ontología a través de una pequeña barra de consultas.

Debajo de la barra de navegación, encontramos las pestañas que vienen por defecto al instalar la aplicación. A continuación, explicamos brevemente su función para más adelante profundizar en cada una de ellas a través de ejemplos:

- **Active Ontology** es la pestaña donde podemos visualizar la información general sobre la ontología y podemos agregar axiomas u otros datos que afectarán a toda

la ontología. Además, podemos agregar anotaciones generales (en forma de *annotation property*).

- En la pestaña **Entities** (figura 41) podemos visualizar todas las entidades OWL 2 (clases, propiedades e instancias) que aparecen en la ontología y la información sobre estas entidades, tanto la que hayamos declarado explícitamente como la que el razonador pueda haber inferido. Además, podemos agregar información sobre estas entidades, por ejemplo, anotaciones en forma de *annotation property*.

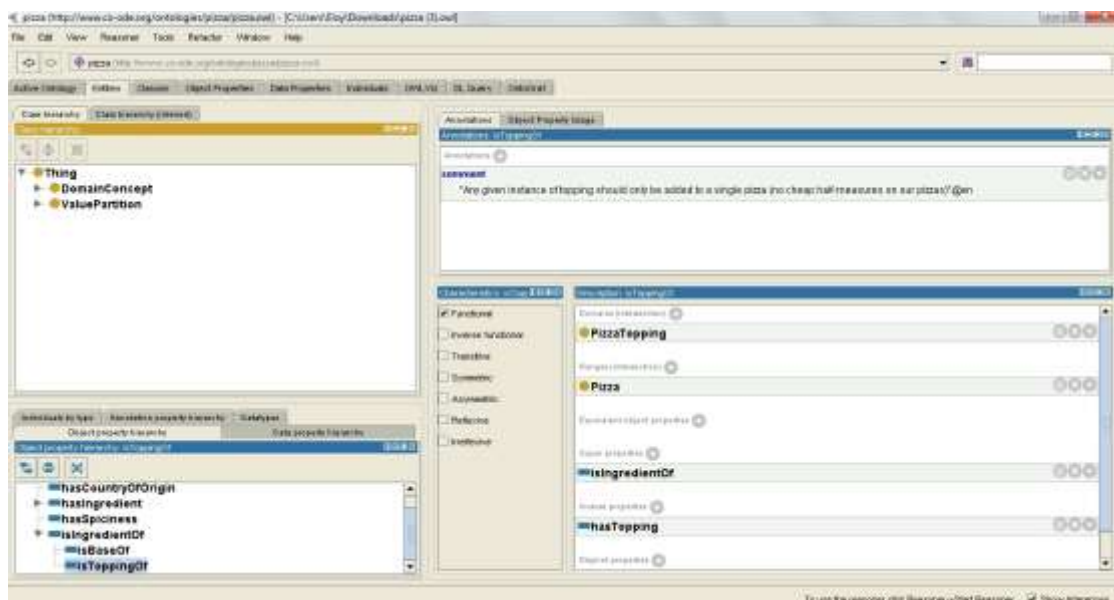


Figura 41. Visualización de la pestaña Entities con la ontología *pizza* abierta

- En la pestaña **Classes** (figura 42) se muestran las clases de la ontología estructuradas de dos formas, la que se haya creado o estemos creando a nivel usuario y la que pueda inferir el razonador. Además, podemos consultar o realizar la descripción de cada una de las clases a través de expresiones con sus propiedades, anotaciones o axiomas.

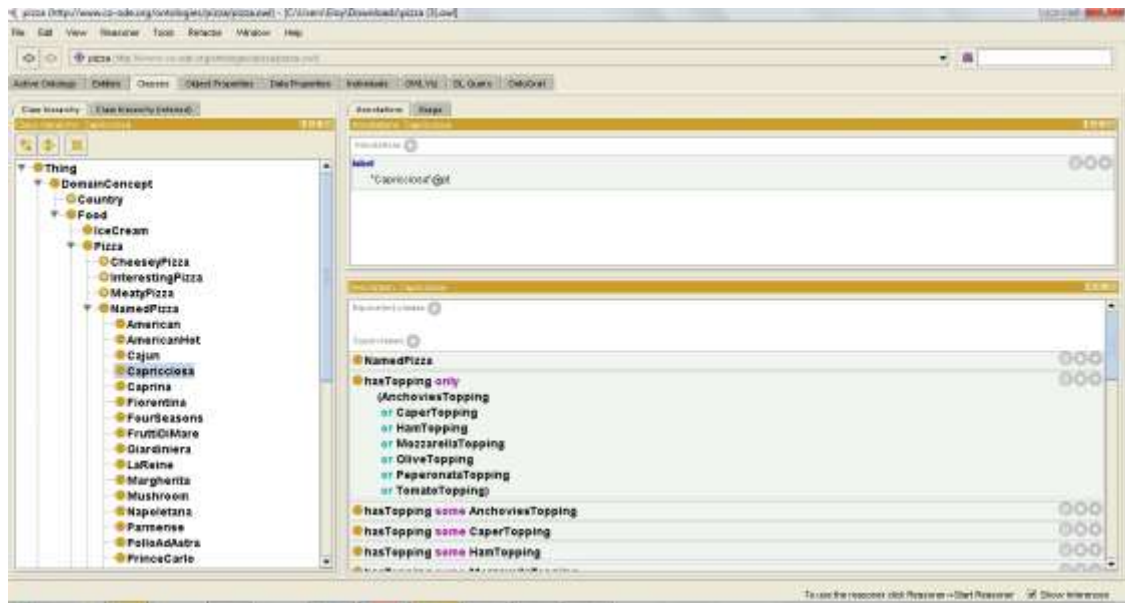


Figura 42. Visualización de la pestaña Classes con la ontología *pizza* abierta

- En la pestaña **Object Properties** (figura 43) visualizaremos y crearemos las propiedades de tipo objeto que formen parte de la ontología. Desde esta pestaña, podemos definir el dominio y rango de las propiedades, indicar sus propiedades matemáticas, incluir anotaciones y organizarlas jerárquicamente.

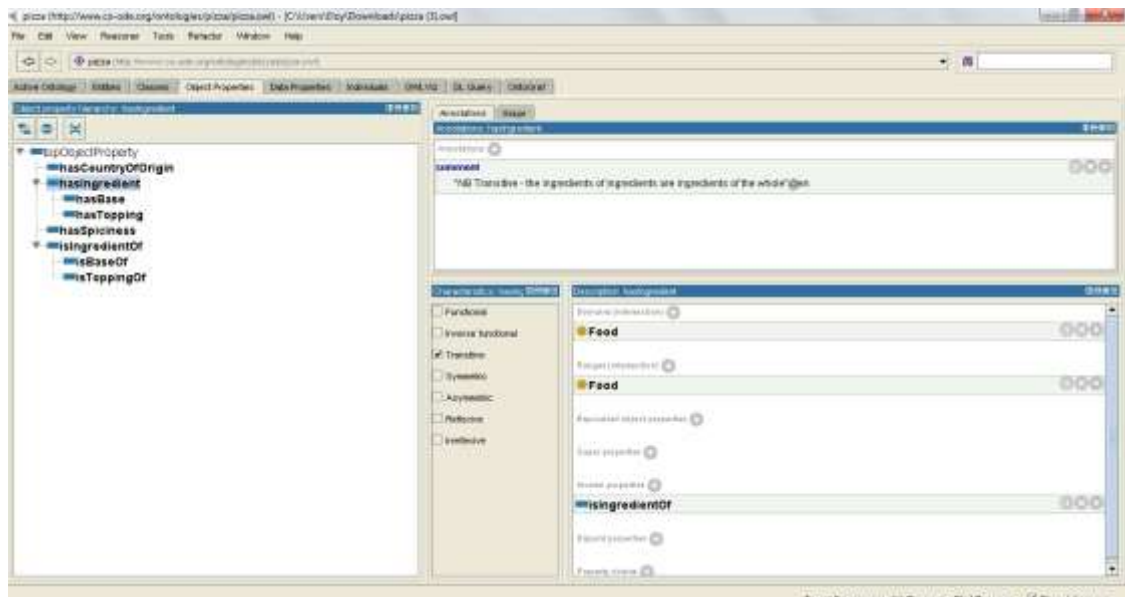


Figura 43. Visualización de la pestaña Object Properties con la ontología *pizza* abierta

- En la pestaña **Data Properties** (figura 44) visualizaremos y crearemos las propiedades de tipo dato que formen parte de la ontología. Igual que en la pestaña anterior, podemos definir el dominio y rango de las propiedades, declarar si es funcional, incluir anotaciones y organizarlas jerárquicamente.

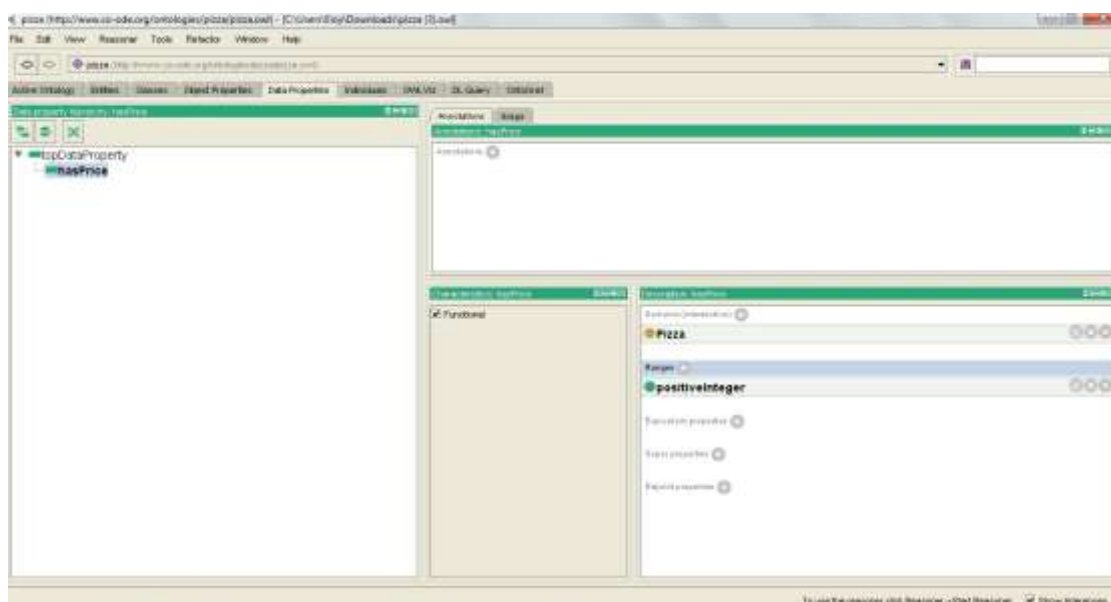


Figura 44. Visualización de la pestaña Data Properties con la ontología *pizza* abierta

- En la pestaña **Individuals** (figura 45) visualizaremos y crearemos las instancias de la ontología, pudiendo anidarlas a clases existentes o describirlas a través de propiedades. Las instancias se mostrarán de dos formas: según las hayamos agrupado al crearlas y agrupadas a clases según el razonador haya inferido.

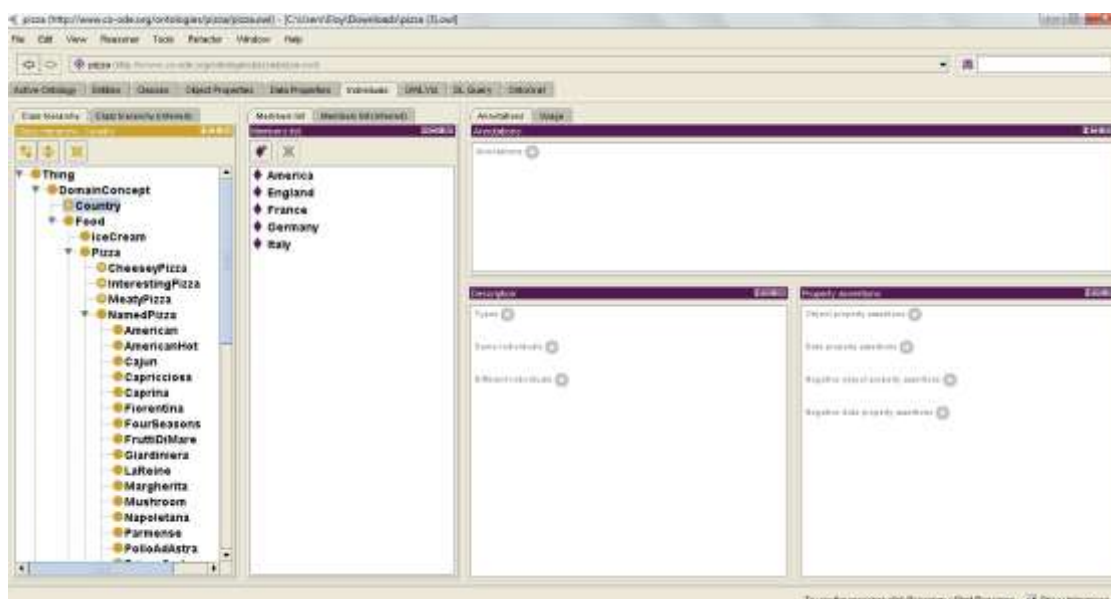


Figura 45. Visualización de la pestaña Individuals con la ontología *pizza* abierta

- Las pestañas **OWL 2 Viz** y **OntoGraf** (figura 46), situadas antes de la pestaña **DL Query**, nos permiten visualizar la ontología de forma gráfica (a través de grafos o esquemas). En el caso de querer utilizar **OWL 2 Viz** necesitaremos tener instalado un software complementario¹⁹.

¹⁹ Software complementario disponible en <http://www.graphviz.org/>

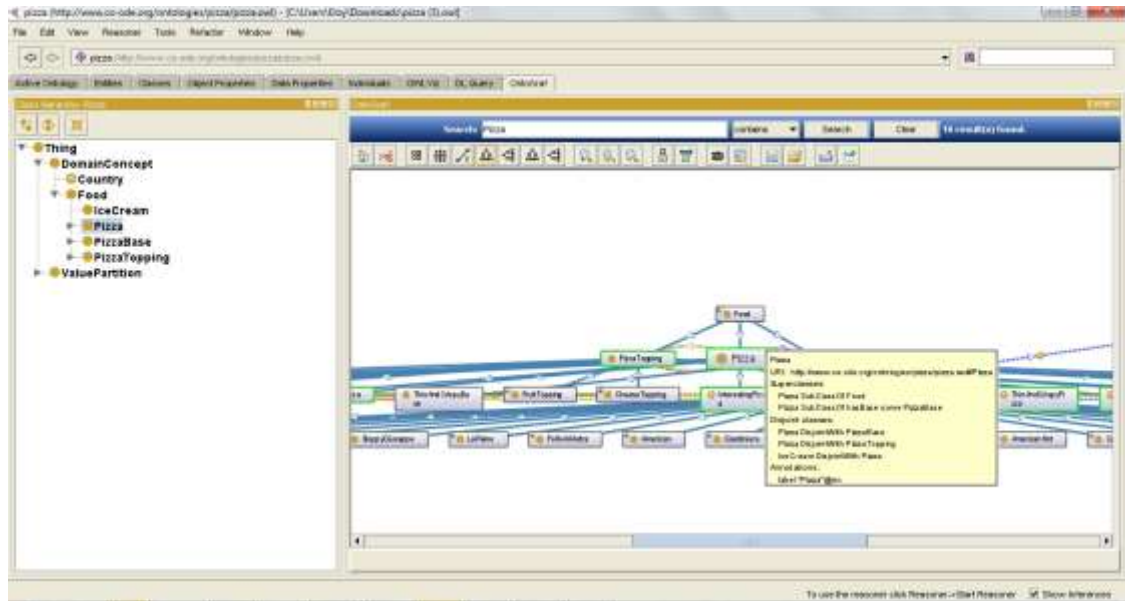


Figura 46. Visualización de la pestaña OntoGraf con el esquema de un fragmento de la ontología *pizza*

- Finalmente, la pestaña **DL Query** (figura 47) nos permite realizar consultas sobre la ontología, estas consultas se basan tanto en la información explícita de la ontología (declarada por el usuario) como en la información implícita que un razonador pueda inferir. Para ello, necesitamos tener activado alguno de los razonadores disponibles para Protégé. En esta pestaña también podemos agregar información nueva a la ontología basándonos en los resultados de las consultas.

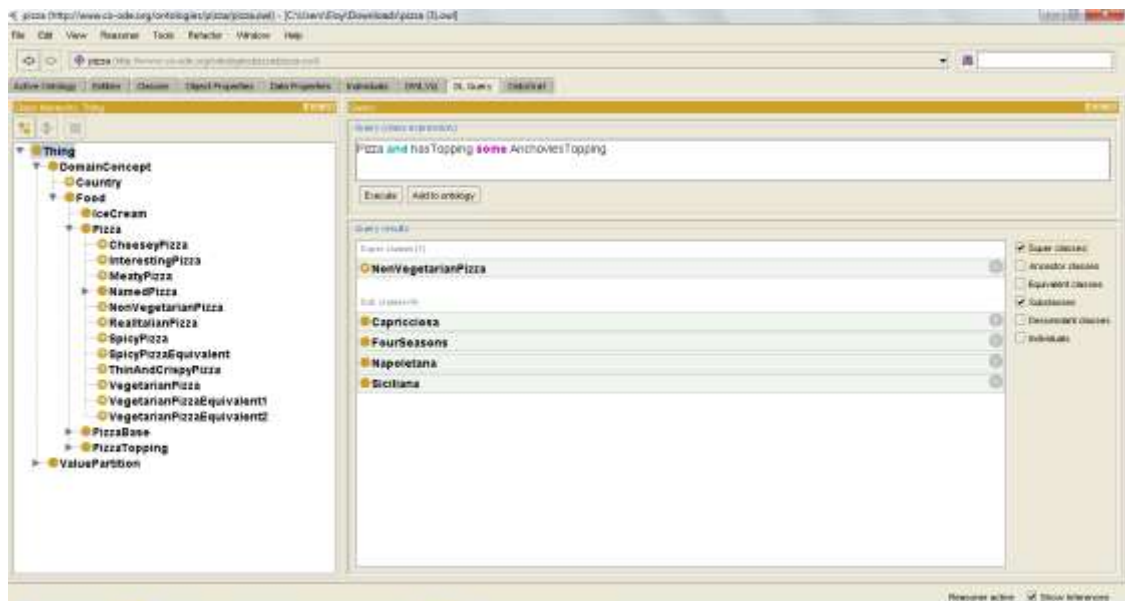


Figura 47. Visualización de la pestaña DL Query con una consulta sobre pizzas que contienen anchoas utilizando el razonador HermiT.

9.1.2 Creación de clases

La creación de las clases de una ontología en Protégé 4 se realiza desde la pestaña **Classes**. Al iniciar la creación de una ontología desde cero observamos que Protégé tiene por defecto una clase denominada *Thing*. Esta clase representa a todas las instancias que puedan

existir en la ontología y la creación de nuevas clases se realiza como subclases de la clase *Thing*. Por ejemplo, en la figura 48, la ontología *Pizzaejemplo* solo tenía la clase *Thing* y en la ventana superpuesta se ha introducido el nombre de una nueva clase, *pizza*, que estará subordinada a la clase *Thing*.

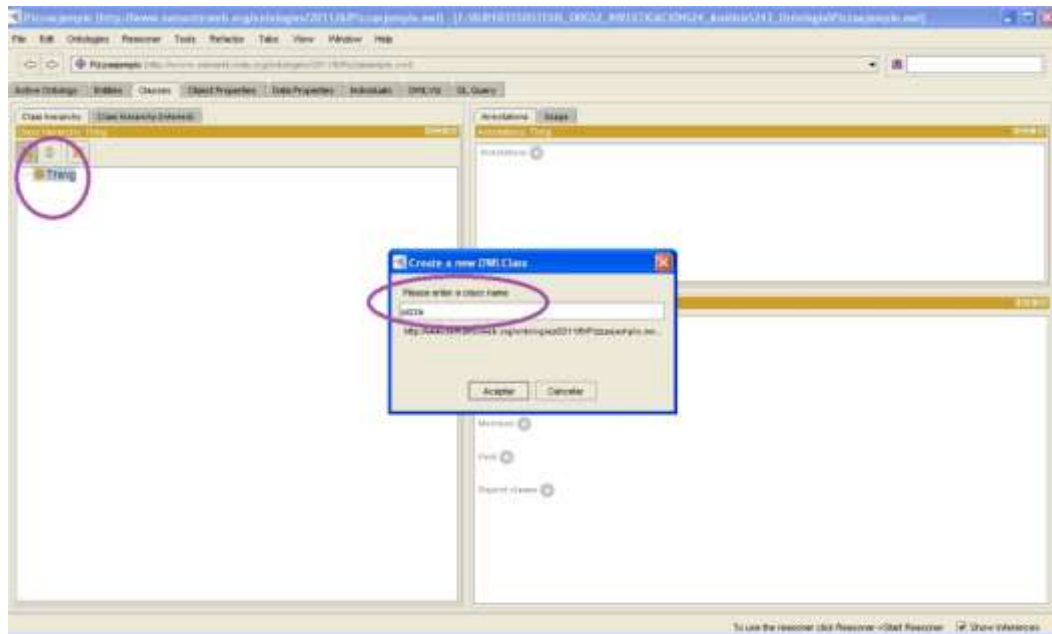


Figura 48. Creación de clase *pizza* bajo la clase *Thing* en la ontología *Pizzaejemplo*

Desde esta pestaña podemos crear toda la jerarquía de clases de la ontología, de modo que unas queden subordinadas a otras según convenga. Por ejemplo, en la figura 49, se observa en el panel izquierdo la jerarquía de conceptos de la ontología *pizza* (Drummond et al., 2007).

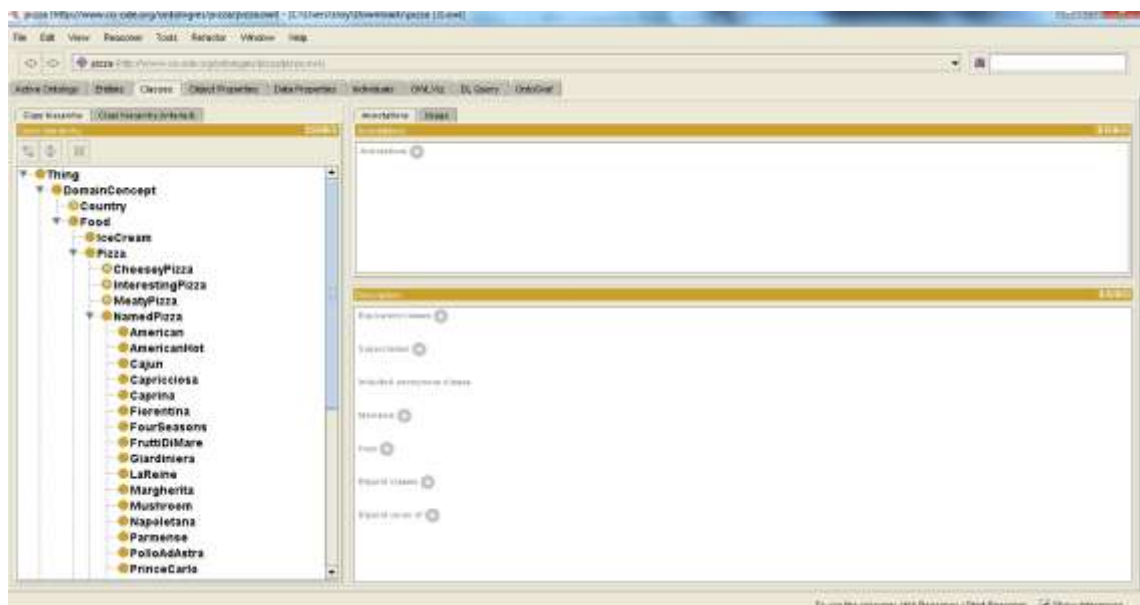


Figura 49. Visualización desde la pestaña *Classes* de Protégé 4 de un fragmento de la jerarquía de *pizza*

En Protégé, no podemos crear dos clases con un mismo nombre, ya que el nombre de una clase es el identificador único de dicha clase. En este caso, podemos entenderlo como que cada clase hace referencia a un concepto y no a una denominación. En el caso de que creamos una

clase con un nombre que ya exista para otra clase, el sistema nos avisará con un mensaje de error y nos impedirá guardar la clase. No obstante, más adelante veremos que existen formas de agregar distintas denominaciones de una misma clase (véase §2.7.1.8).

9.1.3 Creación de propiedades

La creación de propiedades varía según la naturaleza de la relación o característica que queramos representar a través de la propiedad. Principalmente, se distingue entre la creación de una *object Property* y una *datatype Property*. Si queremos crear una *object Property*, nos situaremos en la pestaña correspondiente. En esta pestaña, por defecto, existe una meta-propiedad denominada *topObjectProperty* bajo la cual se crean las propiedades de la ontología. Por ejemplo, en la figura 50, se observa la jerarquía de propiedades de la ontología *Pizzaejemplo*, en concreto estamos visualizado la descripción de la propiedad que representa que una pizza tiene ingredientes. Esta propiedad está representada mediante la *object property tiene_ingrediente* que es una subpropiedad de *topObjectProperty*.

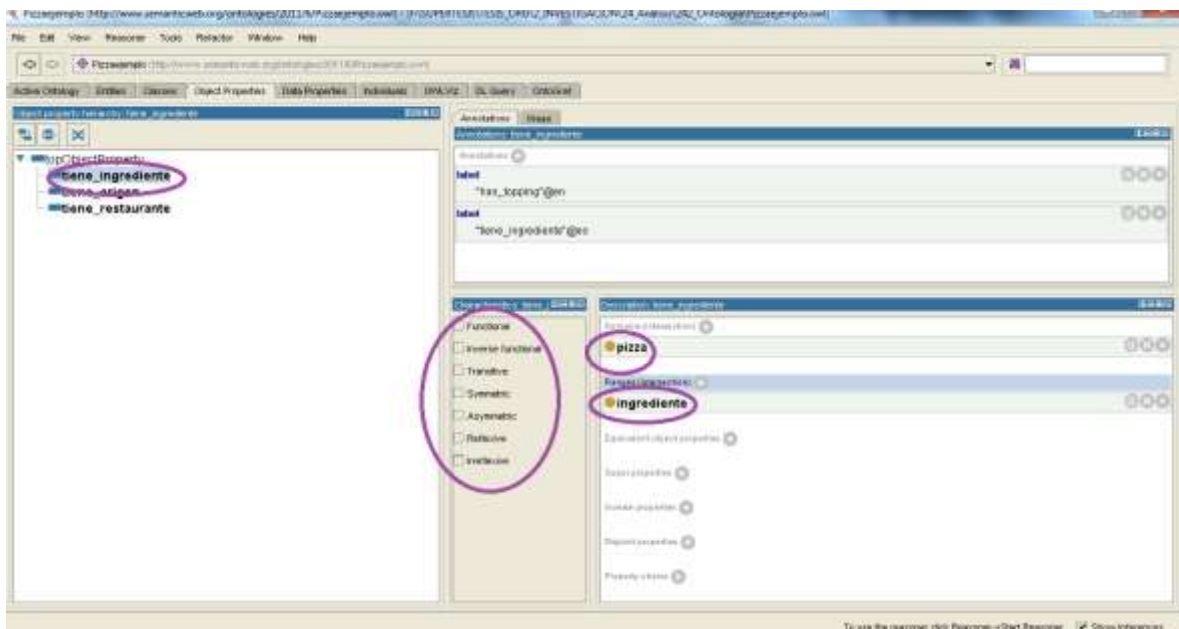


Figura 50. Ejemplo de *object property* en la ontología *Pizzaejemplo*

Si la propiedad fuera transitiva estaría marcada la casilla correspondiente en la subventana *characteristics*. En el caso de *tiene_ingrediente* se han dejado todas las opciones desmarcadas. El dominio y el rango de la propiedad se han declarado en la subventana *description*. En el caso del dominio, en *domain*, está la clase *pizza*, y en el caso del rango, en *range*, está la clase *ingrediente*.

En la descripción de una propiedad, podemos añadir tantos dominios y rangos como sea necesario pero conviene siempre definir como dominio a la clase más genérica que se describa mediante la característica.

Por otra parte, si lo que vamos a representar es una característica o relación cuyos valores no se van a representar en la ontología mediante clases o instancias crearemos una propiedad de tipo *datatype* para representarlos. En Protégé, se visualiza desde la pestaña *Datatype Property* de Protégé y siguiendo la dinámica anterior, sobre *topDatatypeProperty*. Por ejemplo, en la

figura 51 se visualiza la representación de propiedad para expresar el precio de una pizza y se denomina 'tiene_precio'.

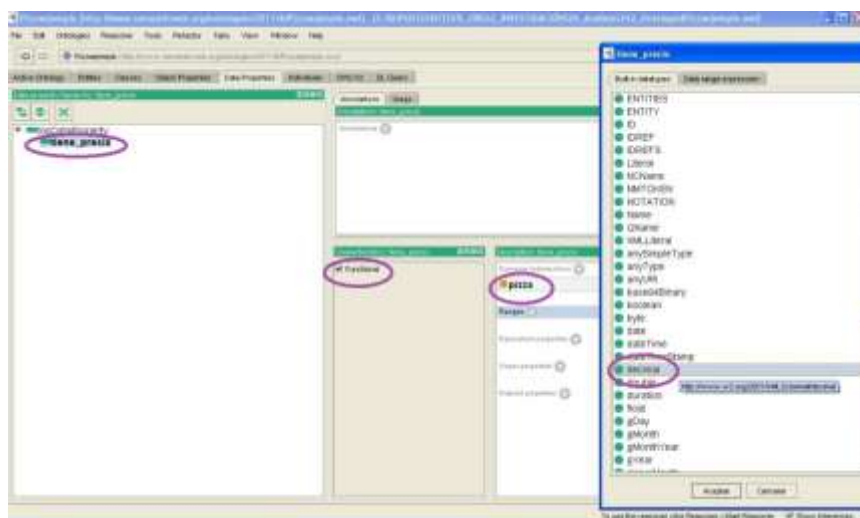


Figura 51. Ejemplo de *datatype property* en *Pizzaejemplo* y ventana de asignación del rango

Además, en la subventana *characteristics* se observa que esta propiedad es de tipo funcional y tiene esta casilla marcada. El dominio de la propiedad aparece en el apartado correspondiente a *domain* y se trata de la clase *pizza*. Los valores de *tiene_precio* son cantidades que incluyen euros y céntimos de euro, para representar esto podemos indicar que el tipo de valor es *decimal* de entre los que se nos ofrecen como tipo de rango.

En Protégé, podemos establecer jerarquías de propiedades, de modo que unas sean superordinadas de otras. Para ello, cuando vayamos a crear la propiedad, en lugar de situarnos sobre *topObjectproperty* o *topDatatypeProperty*, nos situaremos sobre la propiedad que actúe como superordinada. El rango y dominio de la propiedad superordinada es heredado por la propiedad subordinada. Finalmente, conviene señalar que igual que ocurre con las clases, una *object property* no puede, en principio, recibir la misma denominación que otra ya existente y lo mismo ocurre entre propiedades de tipo *datatype*.

9.1.4 Creación de instancias

En Protégé, podemos representar los ejemplares reales de tipo de concepto a través de *individuals* o instancias. Por ejemplo, en el caso de las pizzas, podemos aludir a pizzas reales de una determinada marca o de un restaurante concreto. La creación de una instancia se realiza desde la ventana *individual*.

En la figura 52, se puede ver una instancia que representa la pizza margarita de un restaurante llamado *Etrusco*, a la que hemos denominado *margarita_Etrusco*. Podemos definir bajo qué clase queremos crear la instancia en concreto o realizar una descripción de está de modo que se agrupe bajo una clase que sea acorde con su descripción. En este caso, esta pizza se ha agrupado manualmente bajo la clase *margarita* (subclase de *pizza*) tal y como se ve en el centro inferior de la imagen, en *Types*.

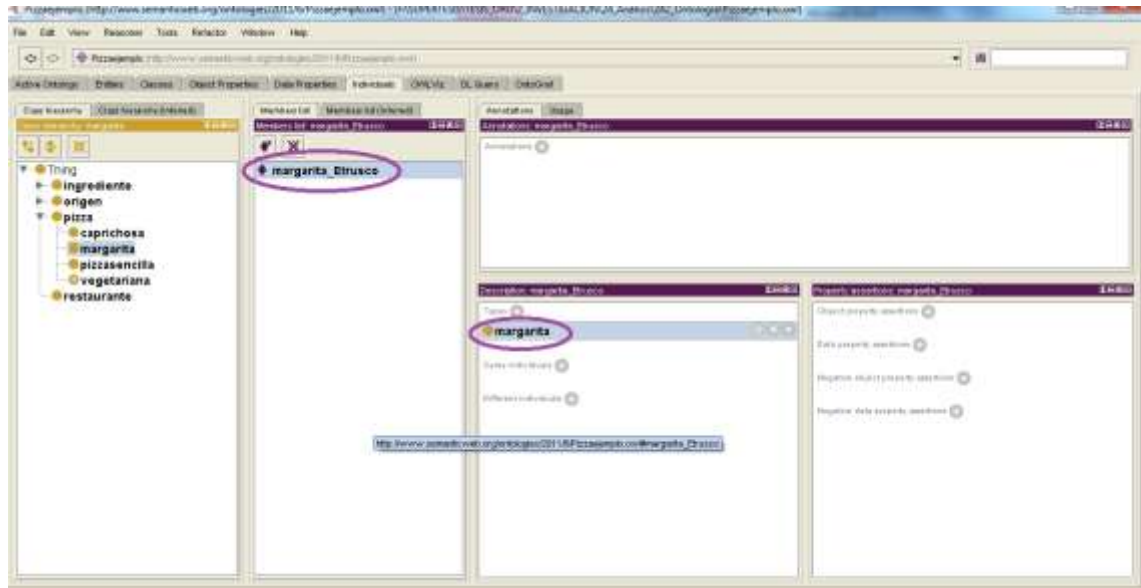


Figura 52. Instancia *margarita_Etrusco* de la clase *margarita* en *Pizzajejemplo*

Igual que el caso de las clases y las propiedades, una instancia no podrá recibir la misma denominación que otra instancia.

9.1.5 Descripción de clases

Para describir las clases de una ontología en Protégé hacemos uso, además de la delimitación del rango y dominio de las propiedades, de la descripción de las mismas clases en la pestaña *classes*.

En esta pestaña encontramos la subventana *description* que contiene, por defecto, 7 apartados. En este caso, nos centraremos, principalmente en los dos primeros, *Equivalent classes* y *Superclasses*, y en el apartado de *Disjoint classes*.

Para que podamos entender mejor estos apartados y su relación con lo que hemos mencionado sobre OWL 2, consideremos que desde *Equivalent classes* podemos agregar expresiones que equivalen a la etiqueta OWL 2 *EquivalentClass*, desde *Superclasses* las expresiones se basan en la etiqueta OWL 2 *SubclassOf* y desde *Disjoint classes* estaríamos declarando una expresión mediante la etiqueta *DisjointWith*.

Asumamos que queremos indicar que una pizza margarita se compone de queso mozzarella y tomate. En Protégé, para declarar esa información, iniciamos la descripción indicando que las clases o instancias que pertenecen a la clase *margarita* tienen esos ingredientes desde el apartado *Superclass*.

Por ejemplo, tal y como se observa en la figura 53, la clase *margarita* contiene en la descripción del apartado *Superclasses* la siguiente expresión: *tiene_ingrediente some tomate*. Así, hemos indicado que una pizza margarita tiene ingredientes y un puede ser tomate. Al mismo tiempo, la expresión *tiene_ingrediente some mozzarella*, indica que una pizza margarita puede tener mozzarella. No obstante, para que una pizza sea margarita no basta con que tenga estos ingredientes sino que es necesario que solo tenga estos ingredientes. En Protégé, esta restricción se realiza indicando que la propiedad solo (*only*) puede tomar esos valores. Si nos fijamos en la figura 53, vemos como la clase margarita *tiene_ingrediente only (tomate and mozzarella)*.

Debemos entender el *and* como una expresión booleana que indica que debe tener ambos ingredientes.



Figura 53. Descripción de la clase *margarita* en *Pizzaejemplo*

A continuación, presentamos un ejemplo para entender el uso del elemento *EquivalentClasses*. Asumamos en nuestra ontología *Pizzaejemplo* hemos agrupado los ingredientes en las siguientes subclases: verdura, carne, pescado, huevo y queso. En la clase de pizzas, queremos tener una pizza que agrupe las pizzas aptas para ovo-lacto-vegetarianos y no queremos reubicarlas en esa clase manualmente sino con la ayuda de un razonador. En Protégé, podemos crear una subclase de pizza, denominada *vegetariana* (figura 54) y desde el apartado *Equivalent classes* podemos describir que la clase *vegetariana* se compone de pizzas que tienen solo ingredientes de la clase verdura, huevo o queso: *pizza and tiene_ingrediente only (verdura or huevo or queso)*.

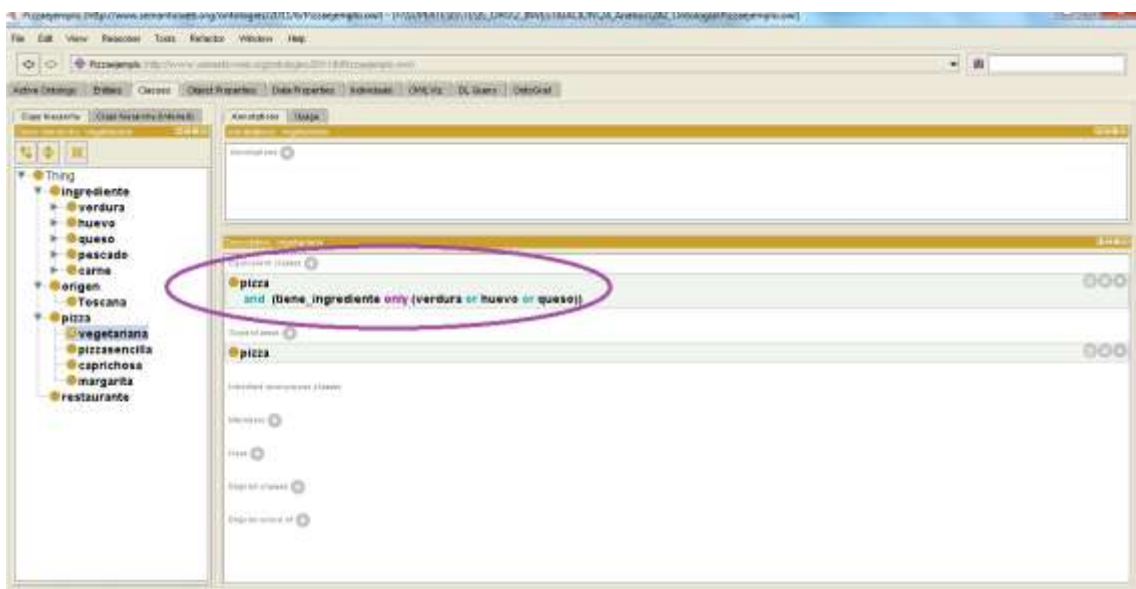


Figura 54. Descripción de la clase *vegetariana* como clase equivalente en *Pizzaejemplo*

De esta forma, si activamos el razonador HerMiT de OWL 2 que viene incluido en la aplicación, en la pestaña de *Class hierarchy (inferred)* que muestra las inferencias que ha realizado el razonador, veremos como la clase *margarita* se ha ubicado automáticamente como subclase de *vegetariana* (figura 55) ya que solo contiene ingredientes del grupo admitido.

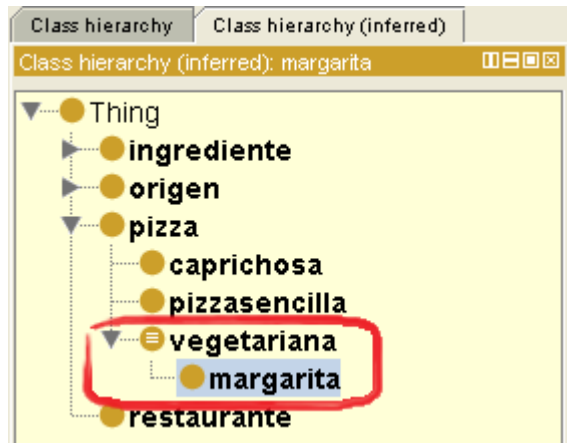


Figura 55. Inferencia de la clase *margarita* como subclase de *vegetariana* en la ontología Pizzatejemplo

Cada vez que creamos una pizza que cumpla estas características se agrupará automáticamente bajo la clase *vegetariana*, independientemente de que pertenezca también a otras clases de pizza.

9.1.6 Descripción de instancias

La descripción de instancias puede realizarse bien a partir de su pertenencia a una clase, tal y como hemos explicado en el apartado de creación de una instancia (§10.1.4) o bien indicando las propiedades específicas que describen la instancia.

Si optamos por la segunda opción, tenemos que acceder a la pestaña *individuals* de Protégé y desde el subapartado *Property assertions* realizar la descripción. Debemos tener en cuenta, que los valores que toma la propiedad de una instancia solo pueden ser instancias o valores correspondientes a algún rango de una *datatype property* (números, cadenas de texto, etc.).

Para realizar la descripción de una instancia necesitamos utilizar una propiedad declarada en la ontología. Por ejemplo, vamos a asumir que tenemos una clase *restaurante* cuyas instancias son los distintos restaurantes donde podemos comer nuestras pizzas (figura 56) y una *property* '*tiene_restaurante*' que vincula las pizzas (dominio) con los restaurantes (rango) donde podemos consumirlas.

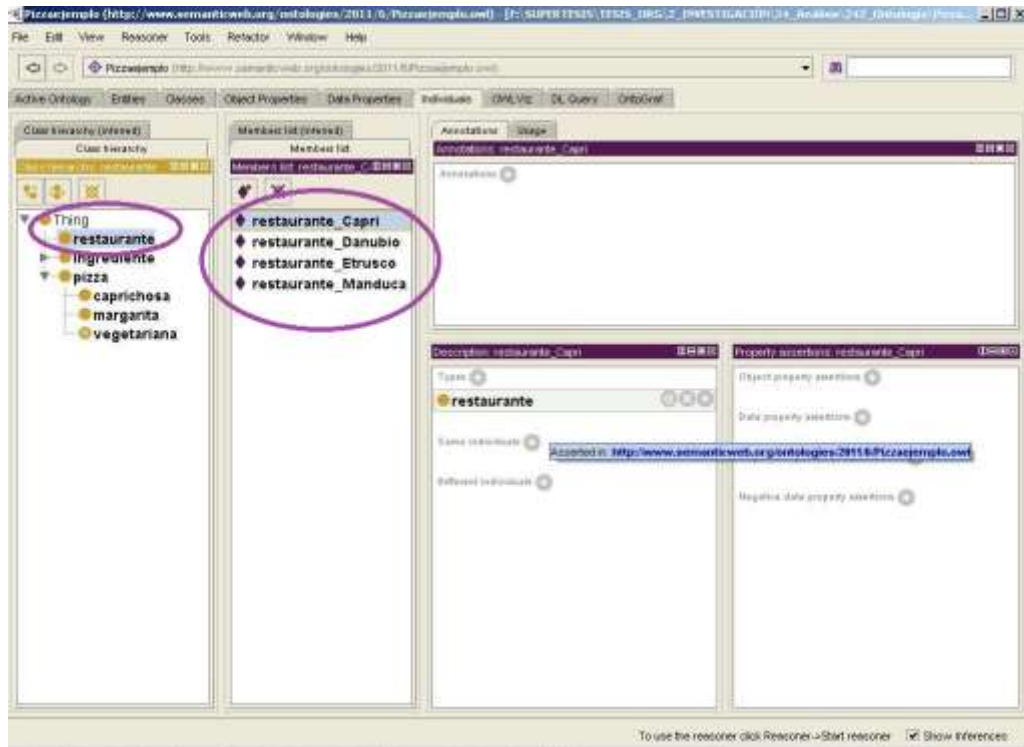


Figura 56. Instancias de la clase *restaurante* en *Pizzaejemplo*

Para describir a la instancia *margarita_Etrusco*, podemos agregar esta propiedad desde el panel *ObjectProperty Assertions* y asignarle un valor concreto. En la figura 57, vemos la instancia *margarita_Etrusco* y cómo en *ObjectProperty Assertions* tiene declarada la propiedad *tiene_restaurante* con valor la instancia *Etrusco* de la clase *restaurante*.

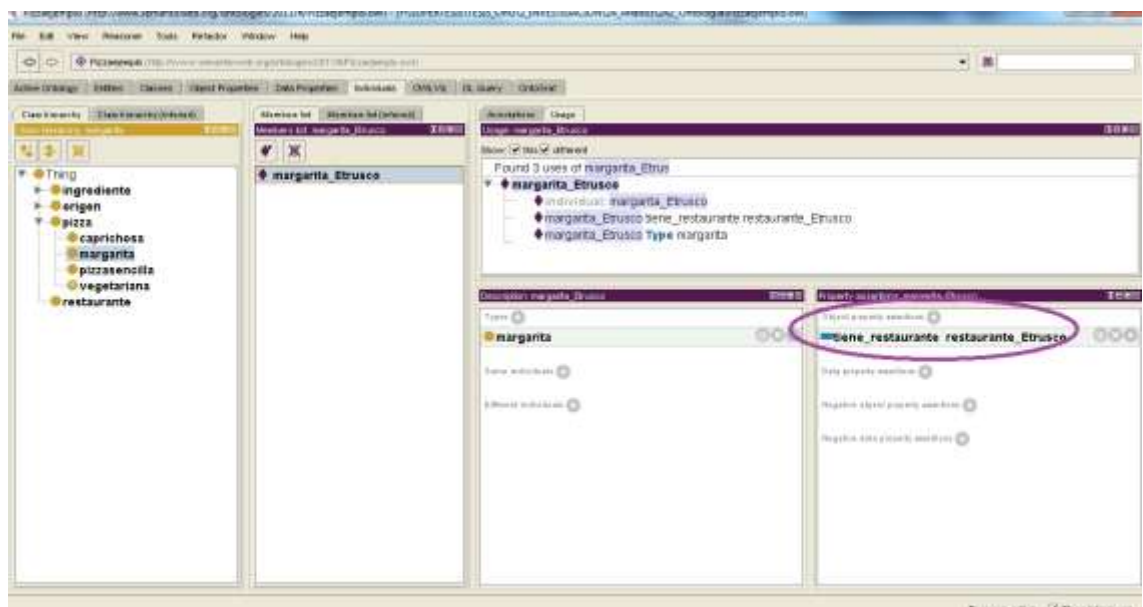


Figura 57. Aserción en la instancia *margarita_Etrusco* de la propiedad *tiene_restaurante* en *PizzaEjemplo*

Otro ejemplo sería asignar un precio a la pizza *margarita_Etrusco*. En este caso como hemos utilizado una *datatype property* para representar que una pizza tiene un precio, la descripción se realiza desde *DatatypeProperty Assertions* en el espacio de *margarita_Etrusco*.

Desde aquí declaramos que, por ejemplo, la pizza margarita del Etrusco vale 6.50, tal y como se observa en la figura 58.

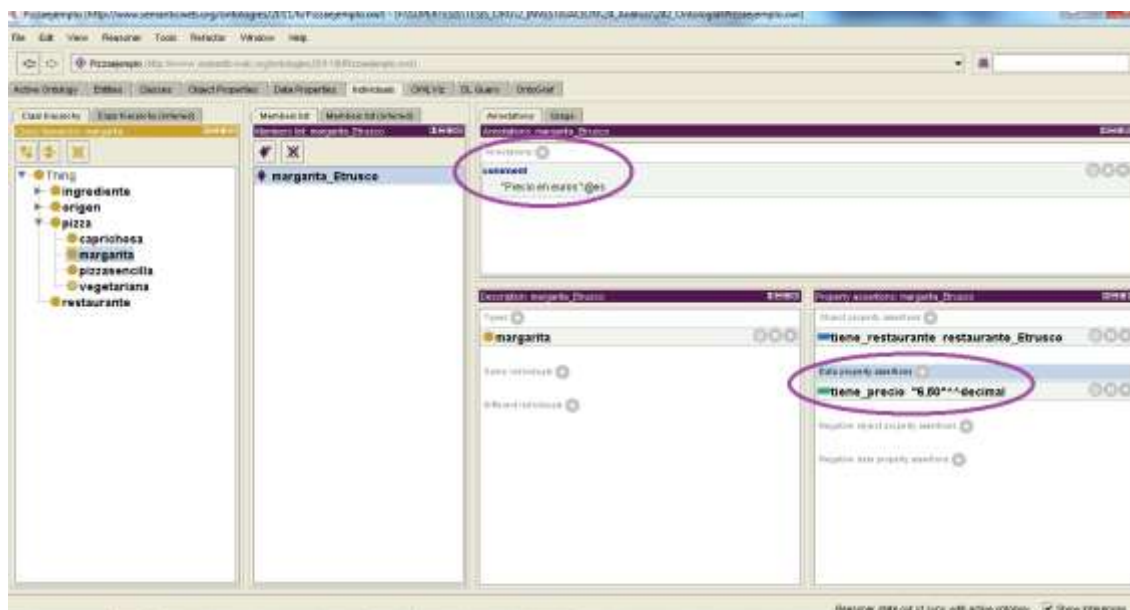


Figura 58. Aserción en la instancia *margarita_Etrusco* de *tiene_precio* y ejemplo de anotación.

Nos puede interesar, además, explicitar que el precio de las pizzas están fijados en euros. Para ello, podemos utilizar una *annotation property* donde indiquemos esta información. Las anotaciones se realizan desde el apartado *Annotations* de la instancia que corresponda.

Como veíamos en OWL 2, los campos que vienen por defecto en una *annotation property* están pensados para diferenciar entre distintos tipos de información. En este caso, hemos declarado que la moneda utilizada son euros, el campo *comment* es un buen candidato para realizar esta declaración. En este tipo de campos podemos o bien escribir en lenguaje OWL 2 o bien en lenguaje natural dependiendo de lo que necesitemos. En el ejemplo de la figura 58 se observa la descripción sobre el tipo de moneda en lenguaje natural. Esta información no será procesada automáticamente por los razonadores como datos a tener en cuenta para establecer las clasificaciones lógicas, pero sí puede recuperarse a través de búsquedas específicas. Por esta razón, en el caso de que nuestra ontología contenga datos en más de una lengua es recomendable indicar el idioma en el que está escrito (en el ejemplo, *es* de español).

9.1.7 Descripción de propiedades

En Protégé, además de describir rasgos sobre la simetría, transitividad etc. de una propiedad OWL 2 también podemos agregar anotaciones que nos ayuden a describir las propiedades en lenguaje natural o a través de enlaces a páginas web u otros recursos.

Esto se realiza desde la pestaña que corresponda al tipo de propiedad, *Object Property* o *Datatype Property*, y una vez allí nos centraremos en el apartado *annotation*.

El apartado de anotaciones funciona del mismo modo en toda la ontología y en las distintas entidades OWL 2 (clases, instancias o propiedades). Existen una serie de campos por defecto que podemos utilizar para describir nuestras entidades o podemos crear nuevos campos para incluir algún tipo de información que consideremos que no encaja en ningún campo existente.

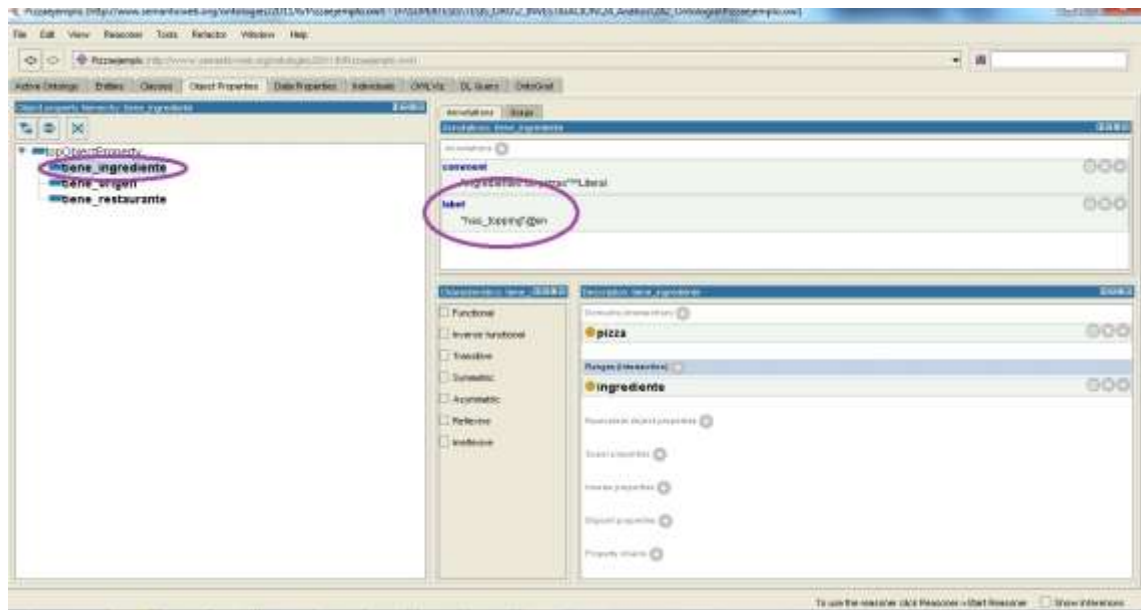


Figura 59. Anotación en la propiedad *tiene_ingrediente* de tipo *label*

Por ejemplo, en la figura 59 se observa una anotación sobre la propiedad ‘tiene_ingrediente’ que indica su denominación en otra lengua. Para ello, hemos utilizado el campo *label* que es el recomendado para incluir denominaciones distintas sobre una misma entidad OWL 2. La denominación se escribe en lenguaje natural, en el ejemplo: “has_topping”. Además, hemos indicado el idioma que hemos usado que se visualiza a continuación del campo de la etiqueta con el símbolo @ seguido del código de idioma. La definición de idioma está disponible para alemán (@de), español (@es), francés (@fr), inglés (@en) y portugués (@pt). La diferenciación entre etiquetas *label* por idiomas permite realizar consultas en la ontología teniendo en cuenta los criterios de idioma, como veremos en el siguiente apartado.

Otro ejemplo, es la anotación en la propiedad *tiene_precio* en forma de comentario (*comment*) que informa sobre el tipo de moneda que se debe usar para asignar el precio (figura 60).

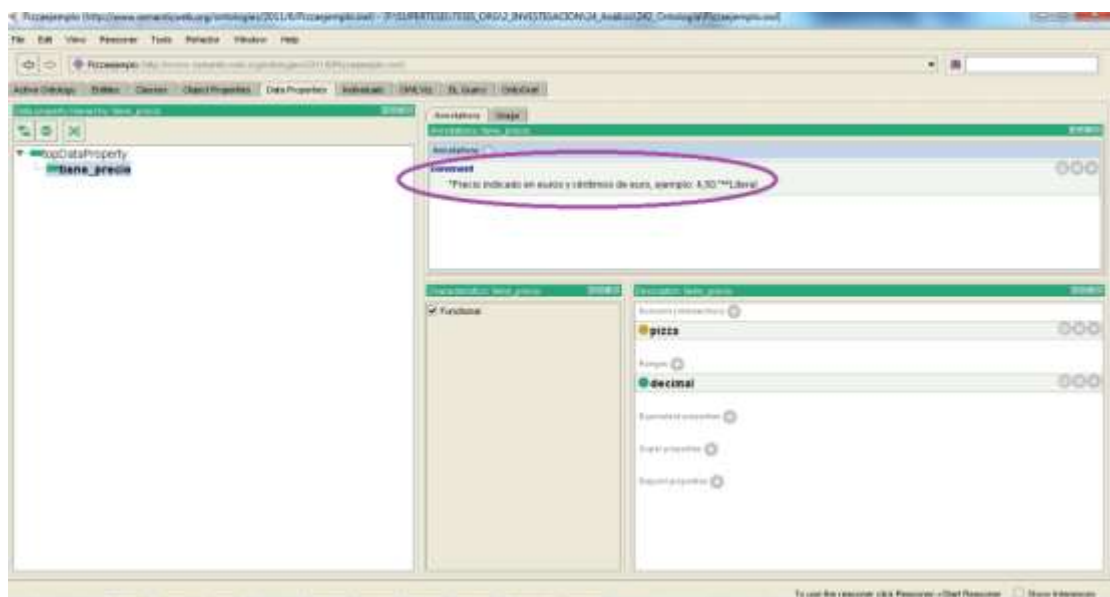


Figura 60. Anotación de la propiedad *tiene_precio* utilizando como tipo *comment*

9.1.8 Multilingüismo y polisemia

Hemos visto como desde Protégé podemos realizar anotaciones e indicar el idioma en que éstas están realizadas. Una forma de crear una ontología multilingüe con Protégé pasa por aprovechar este aspecto y hacer uso de sus potencialidades de visualización. Si, por ejemplo, tenemos la entidad OWL 2 de tipo propiedad ‘tiene_ingrediente’ y la anotamos a través del campo *label* incluyendo su denominación en español y en inglés (figura 61).

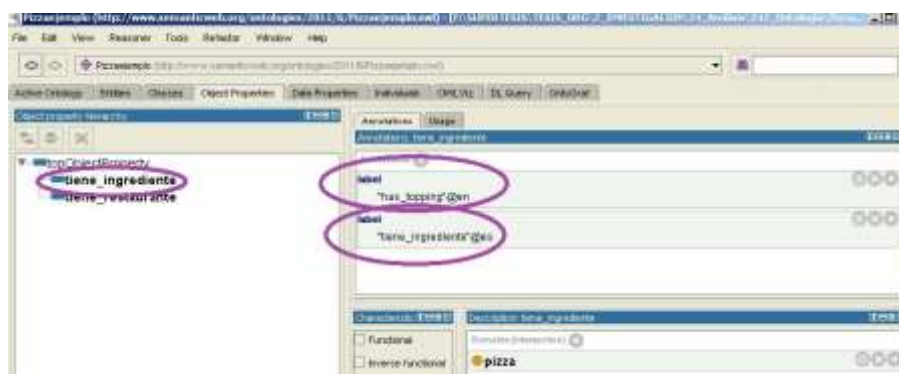


Figura 61. Anotación, mediante *label*, de denominaciones para la propiedad *tiene_ingredientes*

En este punto, la ontología ya podrá visualizarse en una u otra lengua sin hacer uso de ninguna aplicación externa. Para ello, debemos configurar la aplicación Protégé, indicando que deseamos que la ontología se visualice a través del campo *label* y el idioma que nos interese. A continuación mostramos la visualización de las propiedades de tipo objeto a través de *label* en inglés (figura 62).

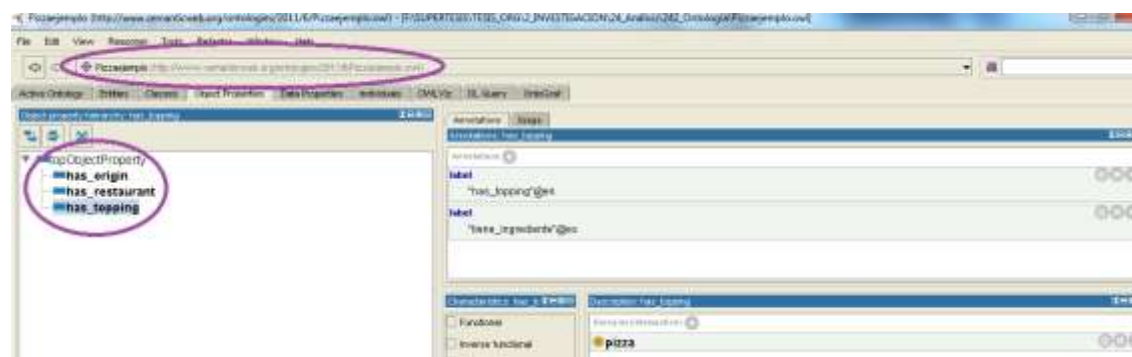


Figura 62. Visualización de propiedades en la ontología *Pizzaejemplo* a través de las etiquetas en idioma inglés.

Una de las principales ventajas es que independientemente de la lengua que tengamos activada, podremos hacer las consultas en la pestaña DL Query en cualquiera de las lenguas de la ontología siempre que exista la denominación que hemos escrito.

Este tipo de anotación de equivalentes a través de *label* se puede realizar en cualquier tipo de entidad OWL 2 (clases, propiedades o instancias).

9.1.9 Creación de axiomas generales

En una ontología puede suceder que queramos realizar una descripción que no afecta directamente a una clase concreta sino a todas las clases que cumplan un determinado requisito. Es decir, a un conjunto de clases que constituyen una clase no explicitada en la ontología (una

clase anónima). La creación de axiomas generales nos permite introducir expresiones sobre estas clases anónimas y describir determinadas naturalezas o funciones en la ontología. Por ejemplo, imaginemos que en nuestra ontología de pizzas queremos describir que todas las pizzas que tengan jamón tendrán como origen Toscana. Esta afirmación se puede declarar en forma de axioma desde la pestaña *Active Ontology*, en la subventana *General class axioms*. Como se observa en la imagen a continuación (figura 45), hemos creado un axioma para recoger el ejemplo, la primera parte del enunciado puede entenderse como “si ocurre x” y la segunda parte que empieza con el *EquivalentTo* se entenderá como un “entonces y”. Este axioma compromete a todos los miembros de la expresión en ambas direcciones, de modo que si una pizza se describe cómo procedente de la Toscana, automáticamente, también se entendería que esta pizza tiene jamón.

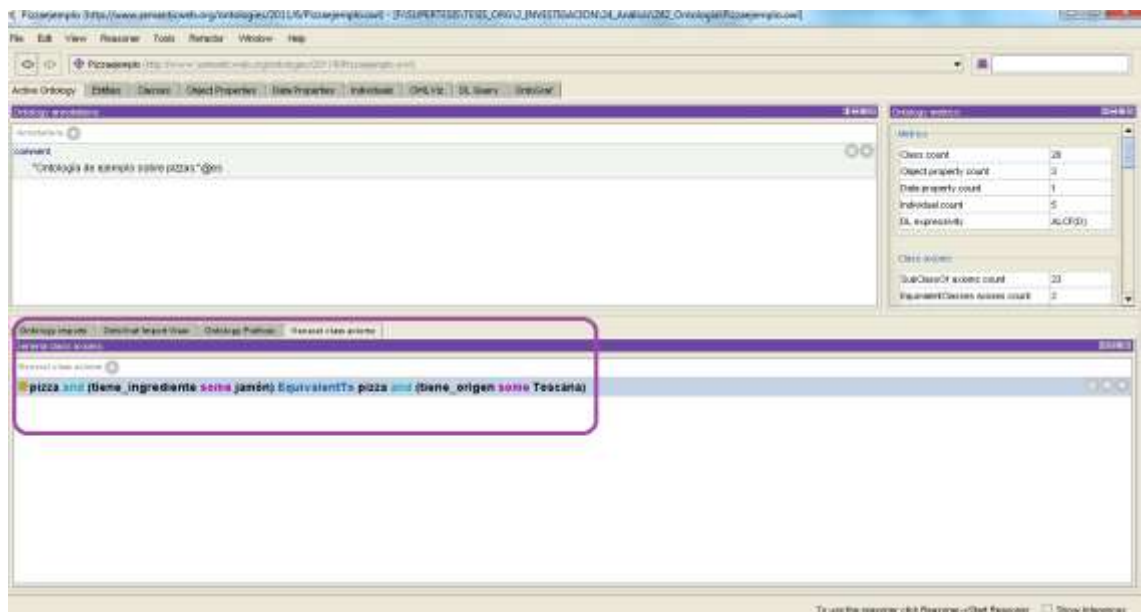


Figura 63. Axioma general en *Pizzaejemplo*

Para observar si el axioma se está considerando, podemos hacer una consulta en la ontología. En la figura 64, se muestra el resultado de la consulta sobre qué pizzas tienen como origen la Toscana y el resultado es la pizza *caprichosa* que no está descrita directamente como procedente de la Toscana pero sí como pizza que contiene jamón.

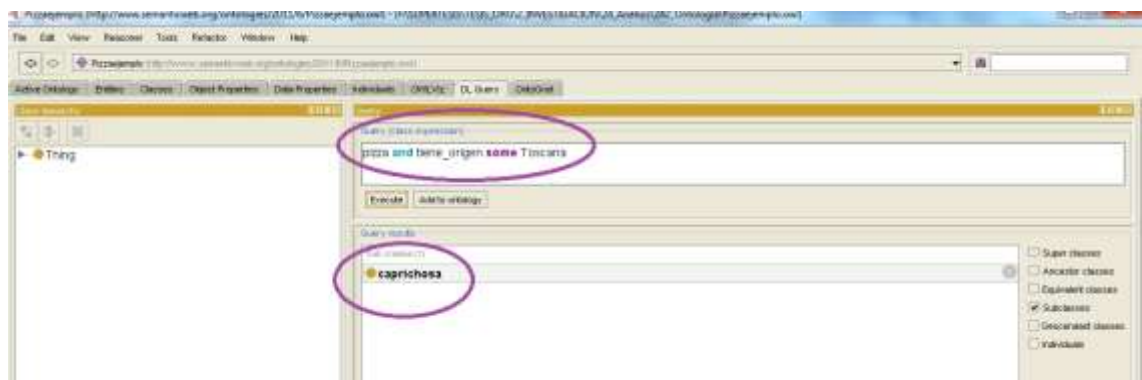


Figura 64. Resultado de la consulta sobre pizzas con origen Toscana en *Pizzaejemplo*

Este tipo de axiomas generales nos pueden servir para introducir las restricciones más generales de un dominio o postulados que se dan por verdaderos en la ontología.

9.1.10 Punning

Antes de finalizar este apartado, conviene que mencionemos una de las novedades del Protégé 4 en relación con las mejoras de OWL 2. Hemos mencionado que en Protégé una clase no puede recibir la misma denominación URI que otra, así como una instancia no puede recibir la misma denominación que otra instancia o una propiedad no puede recibir la misma denominación que otra propiedad. En OWL 2 esto ya es posible pero hay que tener en cuenta que las entidades que compartan denominación URI no siempre se considerarán como una misma entidad por un razonador. Por ejemplo, podemos declarar que existe una propiedad 'ingrediente' y una clase 'ingrediente' y en principio estamos declarado que son una misma entidad con dos facetas distintas (propiedad y clase). Sin embargo, los razonadores más modernos, incluyendo Hermit, todavía no son capaces de procesar esta relación entre clase y propiedad. Por ejemplo, desde el punto de vista del razonador, la información inferida a través de una clase no se adscribirá a la propiedad que reciba la misma denominación.

No obstante, a la espera de que los razonadores puedan procesar estas relaciones entre entidades, es una buena noticia que el lenguaje OWL 2 esté preparado para representar este tipo de información.

9.1.11 Consultas

Como hemos comentado anteriormente, la aplicación Protégé 4 permite hacer consultas sobre la información expresada en la ontología. En la instalación por defecto de la aplicación, estas consultas se pueden realizar desde la ventana *DL Query*. Allí, podemos escribir cualquier expresión sobre la ontología utilizando las propiedades, clases o instancias que la conforman. A continuación, se muestra una consulta sobre tipos de pizza que contienen jamón. Para ello hemos escrito en el apartado de consultas que buscamos pizzas que al mismo tiempo tengan como alguno de sus ingredientes jamón. En la ventana de resultados se muestra la pizza *caprichosa* como una pizza que contiene jamón (figura 65).

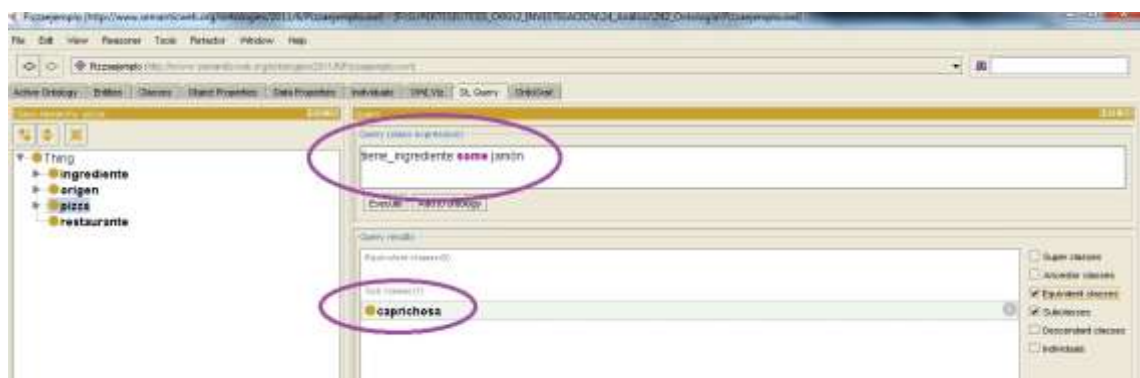


Figura 65. Resultado de la consulta en DL sobre pizzas con jamón

9.1.12 Razonadores para OWL 2 en Protégé

Una razonador de ontologías es un programa que interpreta el lenguaje representado y es capaz de realizar inferencias de información en base al sistema de representación utilizado en el lenguaje. Desde un punto de vista técnico, los razonadores utilizan un motor de inferencia y un conjunto de datos expresados en lenguajes como OWL 2 . Así, permiten realizar inferencias y generar conocimiento a partir de axiomas y hechos declarados en una ontología. Existen distintos algoritmos y métodos de razonamiento (Baader y Nutt, 2003: 78) que se utilizan para

abordar la clasificación automática de conceptos, el análisis de la consistencia de una ontología, la desambiguación de información o la detección de descripciones redundantes. Los resultados que ofrece un razonador con respecto a otro no siempre coinciden (Dentler et al., 2011; Luther et al., 2009; Bock et al., 2008). Por ejemplo, algunos son capaces de interpretar reglas en lenguajes de reglas como SWRL pero otros no, algunos son más rápidos interpretando consultas complejas o clasificando ontologías de gran tamaño que otros, etc.

Las tareas que lleva a cabo un razonador normalmente se dividen en dos grandes grupos (Bock et al., 2008). En un primer grupo encontramos aquellas tareas relacionados con el *TBox* del sistema, donde se almacena la terminología de la ontología (véase §6.3.4), y se describe a través de dos conceptos:

1. *Satisfiability*. Donde se verifica si una clase puede tener alguna instancia de las que forman la ontología teniendo en cuenta la descripción realizada en la ontología.
2. *Subsumption*. Donde se verifica si una clase es subordinada de otra teniendo en cuenta la descripción realizada en la ontología.

A través de estos dos conceptos, se realiza la tarea de clasificación de clases en la ontología, por ejemplo, calculando la jerarquía de subordinación de una ontología.

Las tareas relacionadas con el *ABox*, donde se encuentran las aserciones sobre los individuos mencionados en la terminología del *TBox* (véase §2.2.1.4), normalmente consisten en:

1. Comprobar la consistencia entre lo expresado en el *ABox* y en el *TBox*. Por ejemplo, que las aserciones de una instancia no son incompatibles con las propiedades de la clase a la que pertenece.
2. Comprobar las instancias para ver si cumplen alguna aserción de las expresadas en el *ABox*.
3. Inferir las instancias de una clase, comparando si comparten las mismas propiedades.
4. Comprobar y completar los valores y rangos de las propiedades. Por ejemplo, teniendo una relación *R* y una instancia *i*, recupera todas las instancias *x* que están relacionadas con *i* a través de la propiedad *R*. O recupera el conjunto de propiedades entre dos instancias y comprueba si los valores y rangos que ocupan son adecuados, si coinciden con los rangos y valores que hemos determinado en la ontología.

En la instalación de Protégé 4 por defecto nos encontramos, entre otros, con dos razonadores: Fact++ (Tsarkov y Horrocks, 2006) y HermiT (Shearer et al., 2008). Ambos razonadores han sido desarrollados conjuntamente en la Universidad de Oxford y la Universidad de Manchester y se ofrecen bajo la licencia pública general reducida²⁰ (Lesser GPL en inglés). A continuación, vamos a comparar brevemente estos razonadores y centraremos nuestra descripción no tanto en sus características técnicas sino en el tipo de información que pueden manejar para realizar inferencias y en cómo abordan la detección de inconsistencias.

²⁰ <http://www.gnu.org/licenses/lgpl.html>

Fact++ y HermiT se basan en la lógica de descripciones para realizar las inferencias y el sistema que utilizan para interpretar esta información lógica se basa en lo que se denomina *tableau calculus*. Aunque HermiT va un paso más allá en el desarrollo de su método (*hyper-tableau calculus*), a grandes rasgos podemos pensar que el sistema que utilizan estos programas para clasificar los conceptos y el contenido de una ontología se basa en la búsqueda de modelos que cumplan condiciones consistentes (sin dar lugar a error). A modo de metáfora pensemos que va agrupando las clases según características que tienen en común y las restricciones que hayamos determinado en la ontología y va formando distintos árboles, hasta que determina qué modelo de árbol es el más consistente y a la vez robusto (contiene ramas más pobladas).

Desde un punto de vista aplicado, ambos razonadores son capaces de inferir datos basándose en las propiedades de las clases y en las características de éstas propiedades (simetría, transitividad, etc.). Un aspecto que diferencia a Fact++ de HermiT es que este último soporta más tipos de valor en una datatype. Es decir, Fact++ cuando calcula la estructura de la ontología y su consistencia es capaz de interpretar como valores de propiedad *datatype* los que son de tipo *Literal, string, anyURI, boolean, float, double, integer, int, nonNegativeInteger* o *dateTime*, pero no sería capaz de interpretar las propiedades de tipo *datatype* que tuvieran como valor, por ejemplo, *decimal*. Si nuestra ontología contiene clases descritas mediante una propiedad con un valor de este tipo, al aplicar el razonador nos daría un error. A modo de metáfora, imaginemos que vamos a una tienda de moda en plenas rebajas y solicitamos una prenda en un sistema de tallas que no existe en esa tienda; la dependienta (el razonador), que tiene mucha prisa porque cobra por prenda vendida, nos advertiría que no nos puede ayudar y se pondría con otro cliente.

Esto se debe a que Fact++ todavía no soporta todas las novedades expresivas de OWL 2 y por ello se dice que Fact++ cumple parcialmente con el lenguaje OWL 2 mientras que HermiT cumple completamente las especificaciones de OWL 2.

Cuando se trata de ontologías de gran tamaño, HermiT se ha demostrado que es más rápido y en ocasiones, es capaz de llevar a cabo tareas complejas que otros razonadores no consiguen terminar (Shearer, Motik et al., 2008).

Otra diferencia entre Fact++ y HermiT se encuentra en el uso de reglas en SWRL, que son las siglas de *Semantic Web Rule Language* (Horrocks, Patel-Schneider et al., 2004). El lenguaje SWRL nos permite agregar restricciones a la ontología en forma de reglas. Por ejemplo, si una pizza tiene como ingredientes solo miembros de la clase verduras implica que la pizza es vegetariana. En el apartado de axiomas generales (véase §10.1.9) hemos descrito esta afirmación mediante un axioma general, pero también es posible realizarlo mediante una regla en SWRL. HermiT permite el uso de SWRL en ontologías realizadas en OWL 2 pero Fact++ todavía no tiene implementada completamente esta opción.

9.2 TopBraid Composer

Este editor de ontologías ha sido desarrollado por la empresa privada TopQuadrant²¹ y es un software propietario que ronda los miles de euros, aunque existe una versión de prueba gratuita de la herramienta con opciones más limitadas. Este editor forma parte de un paquete de herramientas denominado *TopBraid Suite* donde el editor se complementa con una herramienta web para relacionar ontologías con aplicaciones web (*Top Braid Ensemble*) y con *Top Braid*

²¹ Se puede consultar la página de TopQuadrant en <http://www.topquadrant.com>

Live que funciona como un razonador y permite recuperar e introducir información en una aplicación basada en ontologías.

El paquete de herramientas está orientado al ámbito de la gestión empresarial y pretende aglutinar y gestionar toda la información relacionada con la arquitectura de una empresa: qué ocurre en la empresa, qué procesos se realizan, qué capacidades tiene, qué sistemas y componentes participan en el funcionamiento de una empresa, qué recursos y tecnología está disponible y cómo se utiliza para conseguir los objetivos de la empresa, cómo se miden los objetivos y la productividad, etc. (TopBraid, 2011).

TopBraid Composer es muy similar a Protégé 4 en apariencia y funcionalidades, en las figuras 66 y 67 vemos la interfaz de TopBraid Composer (licencia Maestro) con una ontología de pizzas cargada y con la descripción de una de las pizzas de la ontología. Una diferencia de apariencia es que las ventanas para crear clases y propiedades no se organizan en forma de pestañas como en Protégé sino que aparecen dispuestas alrededor de la ventana central. De este modo, se nos permite visualizar todo el contenido de la ontología a través de las sub-ventanas que rodean el panel central.

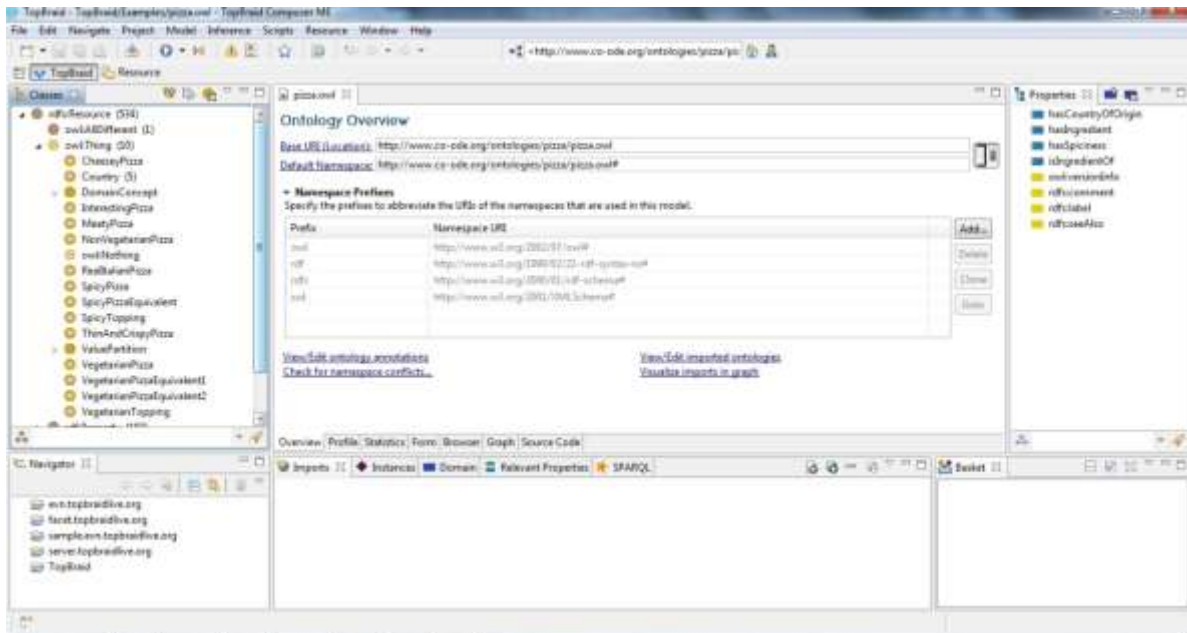


Figura 66. Interfaz de TopBraid con la ontología *pizza ontology*

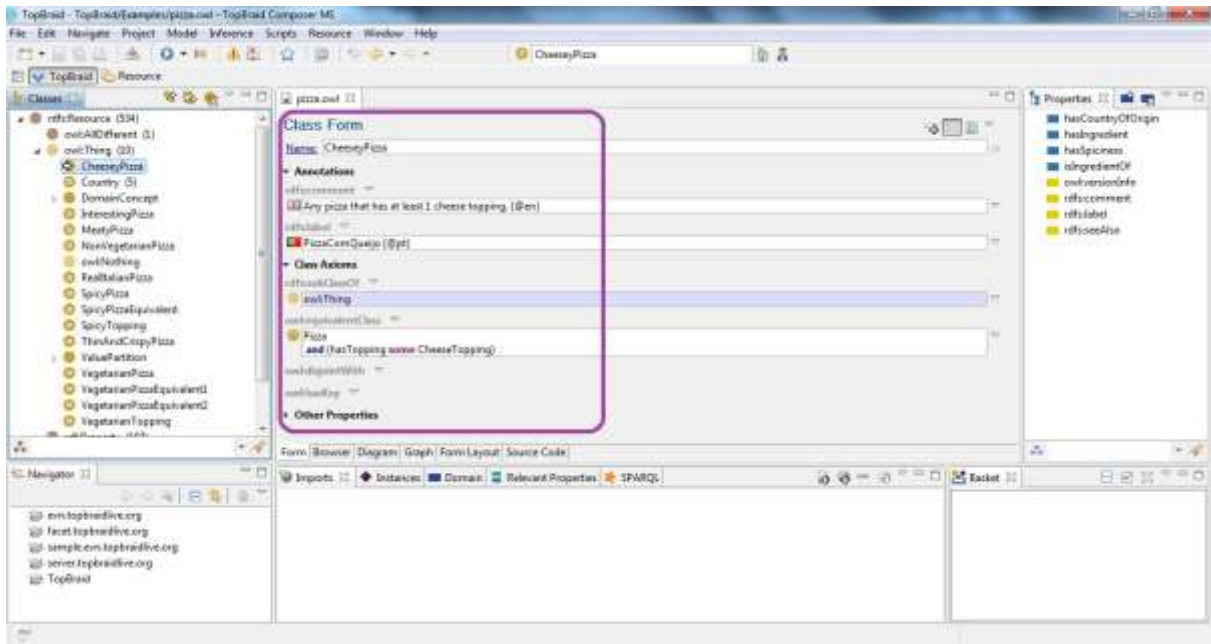


Figura 67. Descripción de la clase *CheesyPizza* en la ventana central

El panel central nos muestra un tipo de información u otra según la actividad que estemos desarrollando. Por ejemplo, si nos situamos sobre una clase de la sub-ventana de clases, nos mostrará la descripción de la clase y si nos situamos sobre una de las propiedades, entonces visualizaremos la descripción de la propiedad.

A simple vista, la aplicación no distingue en su forma de presentar los datos entre propiedades de tipo object y datatype, sino que las muestra juntas en una misma sub-ventana. No obstante, computacionalmente, sí que se realiza la distinción en el momento de crear la propiedad.

La principal ventaja con respecto a Protégé la encontramos en temas de recuperación y etiquetado de contenido. Desde el mismo TopBraid Composer podemos realizar búsquedas del contenido de la ontología en buscadores u otros recursos digitales locales o en línea y recuperar esa información y anotarla para construir y desarrollar la propia ontología.

En la figura 68 vemos el menú de búsqueda y las distintas opciones. Basta con seleccionar una clase o elemento de la ontología y pulsar en alguna de estas opciones para que automáticamente desde la misma interfaz podamos acceder al resultado, por ejemplo, el resultado de la búsqueda de una denominación en Google. En cuanto a la inferencia, TopBraid también integra un razonador de ontologías que, aunque es distinto técnicamente a los que proporciona Protégé, cumple con los mismos objetivos y puede activarse o desactivarse manualmente.

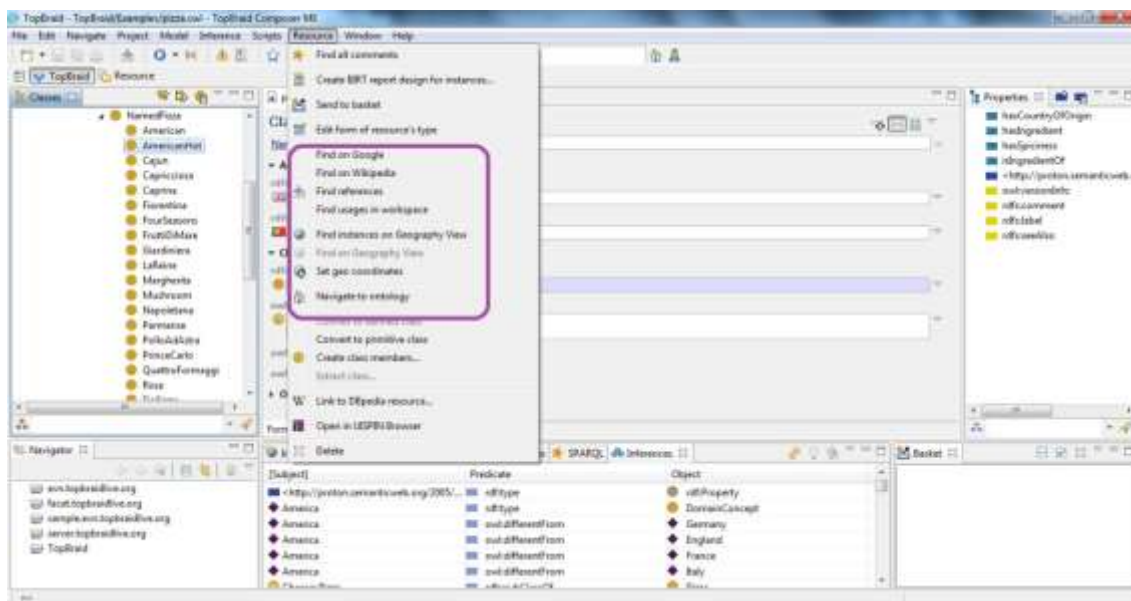


Figura 68. Posibilidades de búsqueda desde TopBraid

Como hemos mencionado, las diferencias a nivel usuario entre ambas aplicaciones son escasas, la forma de crear o describir la ontología también es muy similar a la de Protégé y permite utilizar la sintaxis Manchester OWL 2 syntax.

En lo que respecta al producto generado, TopBraid Composer supone una alternativa compatible con Protégé 4, de modo que podemos editar la ontología con una herramienta y luego abrirla en la otra y seguir editando.

El soporte al producto está orientado a la consultoría privada a través de seminarios no gratuitos y soporte *inhouse*. Aun así, podemos acceder de forma gratuita a algunos documentos informativos a través de su página web²².

9.3 Recapitulación y observaciones

En este apartado hemos revisado distintos tipos de tecnología disponibles para la elaboración de ontologías informáticas, centrándonos especialmente en la herramienta Protégé. Esta herramienta nos permite simplificar el proceso de codificación de conceptos mediante el lenguaje OWL 2 y facilita la elaboración de ontologías. En comparación con TopBraid, las ventajas son su tipo de licencia, la comunidad de usuarios que da soporte a la herramienta y la larga trayectoria y bibliografía disponible sobre la herramienta.

Además del editor de ontologías, hemos visto las potencialidades de los razonadores como herramienta imprescindible para el razonamiento en ontologías.

²² <http://www.topquadrant.com/>

10 Conclusiones generales y síntesis del marco teórico de nuestro trabajo empírico

El objetivo de repasar la bibliografía sobre Terminología e Ingeniería del Conocimiento da como resultado un marco teórico que fundamenta el trabajo empírico que hemos llevado a cabo. A continuación, repasamos las ideas que más nos han influido y debatimos algunas otras con el fin de contribuir a la disciplina de la Terminología con nuestras observaciones.

Los términos se pueden estudiar desde diferentes planos: el lingüístico, el cognitivo y el comunicativo. En el caso de la representación de conceptos, debemos considerar enfoques como el de Roche que nos ayudan a diferenciar entre el plano conceptual y el plano lingüístico de un término. Sin embargo, creemos que también es interesante explorar las posibilidades de un acercamiento a los conceptos a través de las denominaciones y clasificaciones que encontraremos en los textos, siguiendo enfoques como el de Faber. Del plano comunicativo remarcamos la atención a las necesidades del receptor y la descripción de conceptos considerando su situación comunicativa y vinculada a aspectos lingüísticos donde se reflejen los términos de uso. Por eso, nos hemos propuesto estudiar distintas fuentes de un mismo ámbito con el fin de reflejar otras visiones dentro del ámbito e identificar la posible variación de denominaciones dentro del ámbito. En este sentido consideramos igualmente importante contrastar la información recogida con un experto de manera que nuestro trabajo no se base únicamente en una interpretación de los textos.

Consideramos que es necesario un acercamiento híbrido, que combine el acercamiento a los conceptos desde su dimensión lingüística pero también desde su dimensión conceptual

El objeto de estudio de nuestro trabajo son las características de la baldosa cerámica y los conceptos y denominaciones que sirven para describir estas características. Además, nos interesa especialmente su representación en forma de ontología y en este aspecto, coincidimos con Roche al plantear un trabajo donde concepto y término se traten de forma independiente. Enfatizamos la dimensión conceptual al tomar como punto de partida la estructuración de conceptos. Volvemos a coincidir con Roche en la necesidad de construir una ontología que distinga entre una estructura que representa conceptos, y hace uso de términos normalizados para expresarlos en forma de lenguaje, y los términos de uso de cada lengua. El objetivo que perseguimos es que los términos estén relacionados con cada concepto pero sean independientes de la estructura de la ontología.

Nos interesa la estructuración de los conceptos, sean características, relaciones, entidades o procesos, teniendo en cuenta el criterio de diferencia específica pero dejando lugar a la multidimensionalidad. Esto implica la descripción de múltiples diferencias específicas en casos en los que el concepto tiene más de una dimensión en el dominio específico. Por ejemplo, en el caso de ‘baldosa de gres esmaltada’ existen dos diferencias específicas que no ocurren simultáneamente y dependen de los conceptos coordinados del sistema, por un lado, es necesario representar la diferencia entre baldosa de gres esmaltada y baldosa de gres rústica, donde ésta última no es esmaltada pero también diferenciar entre la baldosa de gres esmaltada y el azulejo esmaltado, donde éste último no contiene gres.

Entendemos ‘característica’ como un concepto cuyo referente es una propiedad de otro concepto y consideramos las características en su doble rol de una forma similar a la que

plantean enfoques como el de Madsen o Nuopponen: como elementos que describen los rasgos de un concepto y como elementos que constituyen conceptos en sí mismos. En este aspecto, nos interesa recoger el doble rol cuando realicemos la representación de características, con el objetivo de que la ontología terminológica aporte información sobre las dos facetas.

En la distinción entre relaciones y características consideramos que el enfoque de Madsen et al. (§2.1.4 y 2.1.5) ayuda a entender la semejanza entre una característica y una relación desde el punto de vista de su representación. Entendemos que en una ontología la representación de una relación y una característica (en su rol de descriptor) no difieren en forma ya que ambas suelen utilizar el binomio “atributo:valor”. La distinción se realiza en la etiqueta que utilizemos para denominar el atributo que representa la característica o la relación según consideremos. Por ejemplo, el dominio y el rango de un atributo que representa la relación proceso-paciente será ‘proceso’ y ‘paciente’ (respectivamente), igualmente, el dominio y rango de un atributo que representa la característica ‘tiene paciente’ en un proceso sería ‘proceso’ y ‘paciente’ (respectivamente). En resumen, consideramos que la representación de la relación entre dos conceptos o de la característica depende de los objetivos de representación concretos, de nuestro interés por reflejar el matiz de relación o de cualidad.

En relación con una definición de ontología, entendemos que los acercamientos entre los autores que hablan de bases de conocimiento terminológico y los acercamientos desde la ingeniería del conocimiento a la representación de conceptos en ontologías son muy cercanos.

Entendemos que una ontología terminológica debe basarse en un modelo que se constituye de una ontología formal de conceptos pero a su vez que estos conceptos mantengan un vínculo con la terminología y las relaciones lingüísticas.

En lo que se refiere a la metodología de elaboración de una ontología terminológica coincidimos con Roche y otros autores en que es necesario distinguir entre la fase donde se crea un modelo de conocimiento basado en los fundamentos teóricos de la terminología y la fase en la que este modelo se traslada mediante un lenguaje de formalización a un sistema informático que sirve para representarlo. Y para conseguir este objetivo, consideramos necesario distinguir entre la etapa en la que se estructuran, identifican y delimitan los elementos necesarios para describir las características de una baldosa y la etapa en la que se diseña el modelo que refleja estos elementos en forma de ontología.

Finalmente, en relación con la herramienta de creación de ontologías, a lo largo de la bibliografía, se pueden observar distintos proyectos de desarrollo de ontologías o bases de conocimiento, como CODE4, que desarrollaron aplicaciones *ad hoc*. Lamentablemente, este tipo de aplicaciones dejan de mantenerse al terminar el proyecto y acaban quedando obsoletas para los nuevos sistemas informáticos. En muchos casos, esto impide la reutilización o reciclado de los productos de un proyecto. Dado que perseguimos la reutilización de la ontología, el apoyo a las tecnologías de acceso libre y el uso de un formato de creación estándar, la herramienta Protégé se muestra como la mejor opción para este trabajo. En esta línea, abogamos también por el uso de sintaxis simplificadas, como la Manchester OWL 2 Syntax, que facilitan la creación de ontologías a usuarios poco expertos en el lenguaje OWL 2, o el razonador HermiT, del que existe numerosa documentación y soporte.

III. Descripción del Trabajo Empírico

1 Introducción al trabajo empírico

A continuación, presentamos los resultados del trabajo empírico sobre la conceptualización de las características de la baldosa cerámica en una ontología. El trabajo realizado ha consistido en identificar un conjunto de criterios y elementos necesarios para describir las características. De forma paralela a la fase de identificación de criterios, hemos trabajado en la extracción de la información conceptual necesaria para la descripción de las características de la baldosa cerámica y en recoger una serie de muestras de tipos de baldosa para probar la utilidad de las características. El siguiente paso ha sido elaborar un marco de descripción formal que permitiera representar la información conceptual en forma de ontología y, a continuación, implementar la ontología con la información extraída. El trabajo termina con la evaluación de la consistencia de la ontología y del trabajo realizado.

En cuanto a la estructura de esta sección, en primer lugar, presentamos los recursos y las herramientas (§III.2 y §III.3) que hemos utilizado para llevar a cabo el trabajo. A continuación, en el capítulo 4 (§III.4), presentamos y describimos los criterios y elementos que van a formar parte de la descripción de características. En ese apartado se expone el resultado del análisis realizado para identificar los criterios y elementos que se han utilizado para describir las características, las entidades que están vinculadas y los procesos que formaran parte de la ontología. Los resultados que presentamos proceden del análisis de la bibliografía y de la información obtenida en la fase de extracción.

El capítulo 5 (§III.5) presenta el trabajo de extracción de información y se centra, especialmente, en mostrar la metodología utilizada y los criterios que hemos considerado para extraer, validar y recoger la información conceptual sobre las características de las distintas fuentes. También se exponen los resultados de la validación de los datos por parte de un experto en cerámica y los resultados de esta fase de extracción de información.

El capítulo 6 (§III.6) presenta el marco conceptual formal que estructura la ontología de características. En este apartado se describe con detalle cómo se va a formalizar la información, qué problemas hemos encontrado y cómo los hemos resuelto. Este apartado incluye un resumen del marco utilizado incluyendo información técnica.

El capítulo 7 (§III.7) muestra cómo hemos llevado a cabo la introducción de datos en la ontología. En este apartado se presenta la puesta en práctica de los aspectos planteados en los capítulos 4, 6 y 7 utilizando la información extraída, presentada en el capítulo 5. Finalmente, el capítulo 8 (§III.8) explora la consistencia de la ontología y presenta algunos ejemplos sobre posibilidades de uso.

2 Recursos textuales

A continuación, presentamos los distintos recursos textuales que han servido para extraer información y para realizar consultas sobre el ámbito de la cerámica. Las referencias bibliográficas de estos recursos se pueden consultar en el anexo 10.

2.1 Normas UNE de cerámica

Las normas UNE están desarrolladas por AENOR. AENOR es la entidad española responsable de la elaboración de las normas españolas y representante de los intereses españoles en los organismos de normalización europeos e internacionales.

Tal y como lo define AENOR²³, una norma es un documento de aplicación voluntaria que contiene especificaciones técnicas basadas en los resultados de la experiencia y del desarrollo tecnológico. Las normas se elaboran por comités de expertos y son el fruto del consenso entre todas las partes interesadas e involucradas en la actividad objeto de las mismas. De este modo se pretende que las normas garanticen unos niveles de calidad y seguridad que permitan a cualquier empresa posicionarse mejor en el mercado y constituyen una importante fuente de información para los profesionales de cualquier actividad económica.

En este trabajo hemos utilizado, principalmente, la norma UNE-EN 14411 de 2007 (AENOR, 2007) y la norma UNE-EN ISO 10545 publicada por partes entre 1997 y 2001 como fuentes para extraer el listado inicial de características y alguna información conceptual sobre éstas. Estas normas han sido elaboradas por el Comité Técnico AEN/CNT 138 «Baldosas cerámicas». Este comité está compuesto por expertos del ámbito y colabora con otras entidades de normalización como ISO a través del comité ISO/TC 189 Baldosas cerámicas y el Comité de estandarización europeo CEN/TC 67 Baldosas cerámicas. Para consultas puntuales derivadas de referencias en los recursos utilizados, hemos consultado las normas UNE-ENV 12633 (AENOR, 2003) y UNE-EN 15771 (AENOR, 2011).

2.1.1 Norma UNE-EN 14411

La norma UNE-EN 14411: «Baldosas cerámicas. Definiciones, clasificación, características y marcado.» (AENOR, 2007) define y clasifica las baldosas según algunas de sus características. La norma UNE-EN 14411 es equivalente en contenido y forma a la norma ISO 13006 (ISO, 1998) disponible en alemán, inglés y francés.

La norma UNE-EN 14411 empieza con la definición de 17 conceptos del campo entre los que se incluyen entidades, características y procesos. El primer concepto definido es baldosa cerámica:

«3.1 baldosas cerámicas: Son placas de poco espesor fabricadas con arcillas y/o otras materias primas inorgánicas, generalmente utilizadas como revestimiento de suelos y paredes, moldeadas por extrusión (A) o por prensado (B) a temperatura ambiente, aunque pueden fabricarse mediante otros procedimientos (C), seguidamente secadas y posteriormente cocidas a temperaturas suficientes para desarrollar las propiedades necesarias. Las baldosas pueden ser esmaltadas (GL) o no esmaltadas (UGL) y son incombustibles e inalterables a la luz.

NOTA – Una baldosa totalmente vitrificada es una baldosa con absorción de agua menos del 0,5%.» (AENOR, 2007: 9)

²³ http://www.aenor.es/aenor/normas/normas/quees_norma.asp#.UbSvn_17LSg, fecha de consulta: 09-05-2013.

Además de baldosa cerámica en la norma se definen explícitamente tres subtipos de baldosa: baldosas extruidas, baldosas prensadas en seco y baldosas fabricadas por otros procedimientos.

Los demás conceptos aluden a partes de la baldosa, procesos o instrumentos de clasificación o características. Estas definiciones son muy escuetas, dando por hecho que el lector tiene un conocimiento previo del ámbito. Por ejemplo, la definición de absorción de agua es la siguiente:

«3.9 absorción de agua (*E*): Es el porcentaje de la masa de agua medida según la Norma ISO 10545-3.» (AENOR, 2007: 9)

Por otro lado, la norma incluye tablas de clasificación para las baldosas cerámicas según distintos criterios. En primer lugar, encontramos una clasificación de baldosas según la capacidad de absorción de agua y el método de moldeo. Esta clasificación incluye los valores necesarios para cada grupo de baldosa y los criterios específicos para algunos casos en forma de nota.

Clasificación de las baldosas cerámicas según su absorción de agua y método de moldeo

Moldeo	Grupo I $E \leq 3\%$	Grupo II _a $3\% < E \leq 6\%$	Grupo II _b $6\% < E \leq 10\%$	Grupo III $E > 10\%$
A Extruidas	Grupo AI (véase anexo A)	Grupo AII _{a,1} ^a (véase anexo B)	Grupo AII _{b,1} ^a (véase anexo D)	Grupo AIII (véase anexo F)
		Grupo AII _{a,2} ^a (véase anexo C)	Grupo AII _{b,2} ^a (véase anexo E)	
B Prensadas en seco	Grupo BI _a $E \leq 0,5\%$ (véase anexo G)	Grupo BII _a (véase anexo J)	Grupo BII _b (véase anexo K)	Grupo BIII ^b (véase anexo L)
	Grupo BI _b $0,5\% < E \leq 3\%$ (véase anexo H)			
C Fabricadas por otros métodos	Grupo CI ^c	Grupo CII _a ^c	Grupo CII _b ^c	Grupo CIII ^c

^a Los grupos AII_a y AII_b se subdividen en dos partes (1 y 2) con diferentes especificaciones de producto.
^b El grupo BIII incluye solo a las baldosas esmaltadas. Hay una pequeña cantidad de baldosas prensadas en seco, no esmaltadas con absorción de agua mayor de 10% que no se incluyen en este grupo de producto.
^c Estas baldosas no se incluyen en esta norma europea.

Tabla 16. Tabla de la Norma UNE-EN 14411 sobre tipos de baldosa cerámicas según su absorción de agua y método de moldeo (AENOR, 2007:12)

La tabla 16, muestra los distintos tipos de baldosa cerámica basados en el método de moldeo y los valores de absorción de agua requeridos según el grupo en el que se incluyen. Los datos representados en esta tabla nos permiten extraer terminología sobre nombres de baldosa y características relevantes para éstas. Además, la misma clasificación presente en la tabla nos ayudará en nuestra tarea de representar el conocimiento del ámbito, al presentarnos una clasificación de baldosas y los criterios utilizados para la clasificación, qué tipo de valor tienen las características y qué requisitos son necesarios según la clase de concepto.

En la misma norma, encontramos más detalles sobre la clasificación de baldosas según el método de moldeo y absorción de agua. Para cada grupo la norma presenta una tabla de requisitos que incluye los valores mínimos y máximos que debe tener una baldosa para una serie de características. En la tabla 17, se muestra un extracto de la tabla de requisitos para baldosas de tipo AI donde se determinan los valores mínimos para una serie de características.

Requisitos para baldosas extruidas, Grupo AI, $E \leq 3\%$

Dimensiones y aspecto superficial	Precisión	Natural	Ensayo
Longitud y anchura			
El fabricante debe elegir las dimensiones de fabricación de forma que: a) permitan una junta de anchura nominal entre 3 y 11 mm ^a para baldosas modulares; b) la diferencia entre la medida de fabricación y la medida nominal no sea mayor de ± 3 mm para baldosas no modulares. Desviación admisible en % de la medida media de cada baldosa (2 ó 4 lados) respecto de la dimensión de fabricación (II)	$\pm 1,0\%$ hasta un máximo de ± 2 mm	$\pm 2,0\%$ hasta un máximo de ± 4 mm	ISO 10545-2
Desviación admisible en % de la medida media de cada baldosa (2 ó 4 lados) respecto de la medida media de las 10 probetas (20 ó 40 lados)	$\pm 1,0\%$	$\pm 1,5\%$	ISO 10545-2
Espesor			
a) el fabricante debe especificar el espesor			
b) desviación admisible en % del espesor medio de cada baldosa respecto a la dimensión de fabricación correspondiente.	$\pm 10\%$	$\pm 10\%$	ISO 10545-2
Rectitud de lados^b (cara vista)			
Máxima desviación de rectitud en % con relación a la dimensión de fabricación correspondiente.	$\pm 0,5\%$	$\pm 0,6\%$	ISO 10545-2
Ortogonalidad^b			
Máxima desviación de ortogonalidad en %, con relación a la dimensión de fabricación correspondiente.	$\pm 1,0\%$	$\pm 1,0\%$	ISO 10545-2
Planitud de la superficie			
Máxima desviación admisible de planitud en %:			
a) curvatura central con relación a la diagonal calculada con la dimensión de fabricación	$\pm 0,5\%$	$\pm 1,5\%$	ISO 10545-2
b) curvatura lateral con relación a la dimensión de fabricación correspondiente	$\pm 0,5\%$	$\pm 1,5\%$	ISO 10545-2
c) alabeo con relación a la diagonal calculada con la dimensión de fabricación	$\pm 0,8\%$	$\pm 1,5\%$	ISO 10545-2

Tabla 17. Fragmento de la tabla de requisitos para baldosas extruidas de tipo AI (AENOR, 2007:15).

La tabla 18 muestra otra clasificación de tipos de baldosa distinta a la que presentábamos en la tabla 17. En este caso se especifican los requisitos mínimos para las baldosas según su aplicación. Este tipo de clasificaciones resultan de gran ayuda para construir la ontología ya que delimitan las distintas clases de baldosa y proporcionan información sobre las características que resultan esenciales para describirlas. Asimismo demuestra que en cerámica no existe una clasificación única de baldosas (por comparación con la clasificación presentada más arriba) y nos ayuda a identificar una nueva tipología para ‘baldosa’.

Características	Suelos		Paredes		Método de ensayo
	Interior	Exterior	Interior	Exterior	Referencia
Dimensiones y aspecto superficial					
Longitud y anchura	X	X	X	X	ISO 10545-2
Espesor	X	X	X	X	ISO 10545-2
Rectitud de lados	X	X	X	X	ISO 10545-2
Ortogonalidad	X	X	X	X	ISO 10545-2
Planitud de superficie (curvatura y alabeo)	X	X	X	X	ISO 10545-2
Aspecto superficial	X	X	X	X	ISO 10545-2
Propiedades físicas					
Absorción de agua	X	X	X	X	ISO 10545-3
Resistencia a la flexión	X	X	X	X	ISO 10545-4
Módulo de ruptura	X	X	X	X	ISO 10545-4
Resistencia a la abrasión profunda de baldosas no esmaltadas	X	X			ISO 10545-6
Resistencia a la abrasión superficial de baldosas esmaltadas	X	X			ISO 10545-7
Dilatación térmica lineal ^a	X	X	X	X	ISO 10545-8
Resistencia al choque térmico ^a	X	X	X	X	ISO 10545-9
Resistencia al cuarteo de las baldosas esmaltadas	X	X	X	X	ISO 10545-11
Resistencia a la helada ^b		X		X	ISO 10545-12
Coefficiente de fricción	X	X			Declarar método utilizado
Expansión por humedad ^a	X	X	X	X	ISO 10545-10
Pequeñas diferencias de color ^a	X	X	X	X	ISO 10545-16
Resistencia al impacto ^a	X	X			ISO 10545-5
Propiedades químicas					
Resistencia a las manchas					ISO 10545-14
– baldosas esmaltadas	X	X	X	X	ISO 10545-14
– baldosas no esmaltadas ^a	X	X	X	X	ISO 10545-14
Resistencia a ácidos y álcalis de baja concentración	X	X	X	X	ISO 10545-13
Resistencia a ácidos y álcalis de alta concentración ^a	X	X	X	X	ISO 10545-13
Resistencia a los productos domésticos de limpieza y aditivos para agua de piscinas	X	X	X	X	ISO 10545-13
Emisión de cadmio y plomo de baldosas esmaltadas ^a	X	X	X	X	ISO 10545-15
^a Método disponible, aunque esta norma no especifica valores.					
^b Para baldosas que se destinan a ser utilizadas en lugares donde pueden estar sometidas a condiciones de hielo.					

Tabla 18. Tabla de la norma UNE-EN 14411 sobre características según la aplicación de la baldosa (AENOR, 2007:13)

La norma también incluye información gráfica para ayudar a describir los conceptos expuestos. En la figura 69, se puede observar un extracto de una de las figuras presentes en la norma UNE-EN 14411 que explica qué tipos de medida se tienen en cuenta en una baldosa.

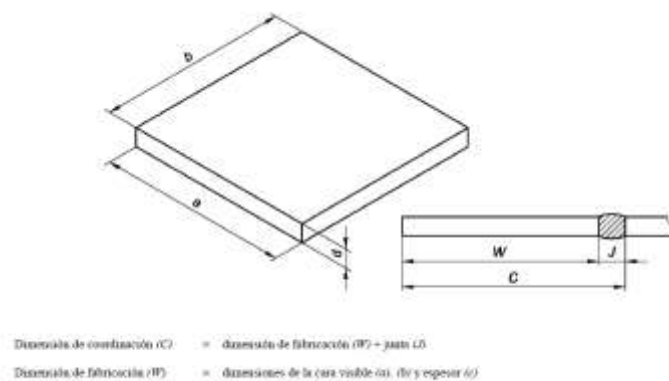


Fig. 1 – Baldosa

Figura 69. Descripción gráfica de baldosa y las distintas medidas que se consideran en la norma (AENOR, 2007:14)

A modo de conclusión, hemos utilizado esta norma por sus posibilidades a la hora de proporcionar listas de conceptos relacionadas con las características del producto acabado. Asimismo, nos ha permitido identificar criterios de clasificación de baldosas basados en algunas características e información de interés sobre cómo distinguir algunos tipos de baldosa y qué características resultan relevantes según la clasificación de baldosas que analicemos.

Esta norma se complementa con la norma UNE-EN ISO 10545.

2.1.2 Norma UNE-EN-ISO 10545

La norma UNE-EN ISO 10545 (AENOR, 1997-2001) es la norma española equivalente a la ISO 10545 (ISO, 1994-2010). Esta norma se compone de 16 partes, cada parte es un documento independiente y de ahí que su publicación se extienda a lo largo de cuatro años.

Cada una de las partes de la norma se centra en una o más características o aspectos relevantes para la descripción de las baldosas cerámicas. En cada parte encontramos información detallada sobre los valores mínimos o máximos permitidos para según qué tipo de baldosa y los procesos relacionados con el ensayo o medida de cada característica. Las 16 partes publicadas hasta 2001 de esta norma son:

1. UNE-EN ISO 10545-1: Baldosas cerámicas. Parte 1: Muestreo y criterios de aceptación (AENOR, 1997a). Equivalente a la norma ISO 10545-1:1995 (ISO, 1995a).
2. UNE-EN ISO 10545-2: Baldosas cerámicas. Parte 2: Determinación de las dimensiones y del aspecto superficial (AENOR, 1998a). Equivalente a la norma ISO 10545-2:1995 (ISO, 1995b).
3. UNE-EN ISO 10545-3: Baldosas cerámicas. Parte 3: Determinación de la absorción de agua, de la porosidad abierta, de la densidad relativa aparente, y de la densidad aparente (AENOR, 1997b). Equivalente a la norma ISO 10545-3:1995 (ISO, 1995c).
4. UNE-EN ISO 10545-4: Baldosas cerámicas. Parte 4: Determinación de la resistencia a la flexión y de la carga de rotura (AENOR, 1997c). Equivalente a la norma ISO 10545-4:1994 (ISO, 1994a).

5. UNE-EN ISO 10545-5: Baldosas cerámicas. Parte 5: Determinación de la resistencia al impacto por medición del coeficiente de restitución (AENOR, 1998^a). Equivalente a la norma ISO 10545-5:1996 (ISO, 1996a).
6. UNE-EN ISO 10545-6: Baldosas cerámicas. Parte 6: Determinación de la resistencia a la abrasión profunda de las baldosas no esmaltadas (AENOR, 1998b). Equivalente a la norma ISO 10545-6:1995 (ISO, 1995d).
7. UNE-EN ISO 10545-7: Baldosas cerámicas. Parte 7: Determinación de la resistencia a la abrasión superficial de las baldosas esmaltadas (AENOR, 1999^a). Equivalente a la norma ISO 10545-7:1996 (ISO, 1996b).
8. UNE-EN ISO 10545-8: Baldosas cerámicas. Parte 8: Determinación de la dilatación térmica lineal (AENOR, 1997d). Equivalente a la norma ISO 10545-8:1994 (ISO, 1994b).
9. UNE-EN ISO 10545-9: Baldosas cerámicas. Parte 9: Determinación de la resistencia al choque térmico (AENOR, 1999b). Equivalente a la norma ISO 10545-9:1994 (ISO, 1994c).
10. UNE-EN ISO 10545-10: Baldosas cerámicas. Parte 10: Determinación de la dilatación por humedad (AENOR, 1997e). Equivalente a la norma ISO 10545-10:1995 (ISO, 1995e).
11. UNE-EN ISO 10545-11: Baldosas cerámicas. Parte 11: Determinación de la resistencia al cuarteo de baldosas esmaltadas (AENOR, 1997f). Equivalente a la norma ISO 10545-11:1994 (ISO, 1994d).
12. UNE-EN ISO 10545-12: Baldosas cerámicas. Parte 12: Determinación de la resistencia a la helada (AENOR, 1997g). Equivalente a la norma ISO 10545-12:1995 (ISO, 1995f).
13. UNE-EN ISO 10545-13: Baldosas cerámicas. Parte 13: Determinación de la resistencia química (AENOR, 1998c). Equivalente a la norma ISO 10545-13:1995 (ISO, 1995g).
14. UNE-EN ISO 10545-14: Baldosas cerámicas. Parte 14: Determinación de la resistencia a las manchas (AENOR, 1998d). Equivalente a la norma ISO 10545-14:1995 (ISO, 1995h).
15. UNE-EN ISO 10545-15: Baldosas cerámicas. Parte 15: Determinación de la emisión de plomo y cadmio en las baldosas esmaltadas (AENOR, 1998e). Equivalente a la norma ISO 10545-15:1995 (ISO, 1995i).
16. UNE-EN ISO 10545-16: Baldosas cerámicas. Parte 16: Determinación de pequeñas diferencias de color (AENOR, 2001). Equivalente a la norma ISO 10545-16:1999 (ISO, 1999).

Cada parte de la norma ISO 10545 se estructura por apartados. Los apartados más comunes son objeto y campo de aplicación, fundamento, equipo, probetas, procedimiento, expresión de resultados e informe de ensayo.

El apartado de «Objeto y campo de aplicación» explica la motivación de la norma. Puede incluir información sobre la relevancia de la característica o aspectos tratados en la norma para la clasificación de baldosas cerámicas, sobre los límites de la norma o sobre procesos específicos vinculados a la norma.

El apartado «Fundamento» indica la característica o características en que se centra la norma centrándose especialmente en el proceso o metodología que se va a aplicar para identificar los valores de la característica o características. No se trata de una definición formal y a menudo, la explicación sólo hace referencia al proceso o los criterios que deben considerarse al calcular los valores de una característica en las baldosas. Por ejemplo, todo el apartado «Fundamento» de la norma UNE-EN ISO 10545-1 (AENOR, 1997a) se compone del siguiente texto:

«Esta parte de la Norma ISO 10545 regula un sistema de inspección por muestreo doble, en parte por el método de inspección por atributos (valores individuales), y en parte por un método de inspección de valores medios (variables).

El número de baldosas a ensayar varía según la propiedad considerada (véase tabla 1).»

De este texto se desprende que la motivación es identificar el método de trabajo para determinar los valores a ensayar pero no se incluye información sobre qué se entiende por muestreo sino que la mayor parte de la información se asume como conocida por el lector.

El apartado «Equipo» especifica las herramientas o aparatos que se van a necesitar para llevar a cabo el ensayo que describe la norma. En este apartado encontramos listas concisas de aparatos y algunas especificaciones al respecto. Por ejemplo, en la norma UNE-EN ISO 10545-3 sobre determinación de la absorción de agua, de la porosidad abierta, de la densidad relativa aparente y de la densidad aparente, en el apartado de «Equipo», incluye una lista de 10 elementos entre aparatos y materiales como «3.1 Estufa de secado capaz de funcionar a (110 ± 5) °C» o «3.6 Desecador». De nuevo, esta información hace referencia a conceptos que no aparecen descritos en la norma, por ejemplo qué es una ‘estufa de secado’.

El apartado «Probetas» incluye detalles sobre el tipo de baldosas que se van a considerar en el ensayo y sobre las especificaciones de las muestras de baldosas que se pueden ensayar. En este apartado podemos encontrar denominaciones para distintos tipos de baldosa.

El apartado «Procedimiento» explica cómo evaluar o medir los valores de la característica a ensayar en las baldosas. Este apartado se divide, a menudo, en subapartados según las distintas fases y los distintos métodos disponibles. En este apartado podemos encontrar información conceptual relevante sobre los procesos que están vinculados con las características. Por ejemplo, siguiendo con la norma UNE-EN 10545-3 (AENOR, 1997b), se mencionan dos fases para ensayar la absorción de agua en las baldosas: ‘impregnación de agua’ y ‘pesaje hidrostático’, y en el caso de ‘impregnación de agua’, se mencionan dos métodos para obtener el resultado: ‘método por ebullición’ y ‘método al vacío’.

El apartado «Expresión de resultados» incluye información detallada sobre cómo expresar los resultados del ensayo en baldosas de la característica o características tratadas en la norma. Puede incluir fórmulas o tablas de valores y suele indicar las unidades de medida en las que se expresarán los valores. Este apartado también puede incluir clasificaciones de baldosas basadas en los criterios de ensayo que trata la norma. Por ejemplo, siguiendo con el contenido de la norma UNE-EN ISO 10545-3, el apartado «Expresión de resultados» está subdividido, en este caso, en distintos sub-apartados sobre cómo expresar los resultados para cada característica tratada en esta parte de la norma: absorción de agua, porosidad abierta, densidad relativa aparente y densidad aparente. En el apartado de «absorción de agua» encontramos la siguiente información (AENOR, 1997b:6):

6.1 Absorción de agua

Para cada baldosa, el coeficiente de absorción de agua, $E_{(p,v)}$, expresado como porcentaje de masa seca, se calcula mediante la ecuación:

$$E_{(p,v)} = \frac{m_{2(p,v)} - m_1}{m_1} \times 100$$

donde

m_1 es la masa de la baldosa seca;

m_2 es la masa de la baldosa húmeda.

La designación E_p debe emplearse para el coeficiente calculado mediante m_{2p} y E_v , cuando se emplea m_{2v} . E_p representa la penetración del agua en los poros de fácil impregnación, mientras que E_v representa la penetración del agua en la mayor parte de los poros abiertos.

De este fragmento podemos extraer que la ‘absorción de agua’ en las baldosas se expresa a través de un número seguido por un porcentaje. A lo largo de toda las partes de la norma UNE-EN ISO 10545 vamos a encontrar en los apartados de «Expresión de resultados» información sobre cómo se expresan los valores de las características analizadas lo que resulta de gran utilidad para nuestro trabajo.

Otro ejemplo que forma parte de la UNE-EN ISO 10545-7 (AENOR, 1999a), sobre resistencia a la abrasión profunda en baldosas esmaltadas, establece un número de clase a las baldosas según el resultado del ensayo (tabla 19):

Clasificación de baldosas cerámicas esmaltadas

Etapa de abrasión; defecto visible a revoluciones	Clase
100	0
150	1
600	2
750, 1 500	3
2 100, 6 000, 12 000	4
>12 000 ¹⁾	5

1) Debe superar el ensayo especificado en ISO 10545-14 para la resistencia a las manchas.

Tabla 19. Clasificación de baldosas cerámicas esmaltadas según su resistencia a la abrasión profunda en la norma UNE-EN 10545-7 (AENOR, 1999a: 9)

A partir de este fragmento podemos aprender que la característica ‘resistencia a la abrasión profunda’ en baldosas esmaltadas tendrá como valores posibles un número entre 0 y 5.

Finalmente, mencionar que las distintas partes de las normas suelen incluir un último apartado sobre cómo hacer el informe del ensayo realizado. Este tipo de apartados hace referencia a la documentación que hay que elaborar para justificar los resultados del ensayo.

En general, atendiendo a la terminología y a la información terminológica presente en las distintas partes de la norma, este tipo de documentos nos ha resultado útil para identificar denominaciones para las características de la baldosa e información sobre su significado. La información relativa a cómo se expresan los valores de las características ha resultado esencial para completar nuestra ontología Hemos encontrado además información sobre los procesos relacionados con la medición o ensayo de las características y las herramientas utilizadas que nos han ayudado a completar la descripción de las características. Finalmente, el contenido y análisis de estos documentos también ha contribuido en la identificación de las características

esenciales para los distintos tipos de baldosa y los criterios para clasificar las baldosas de acuerdo con los valores de las características presentes en esta norma y la norma que hemos presentado anteriormente.

Las características de formato de estas normas, formato pdf protegido, han impedido el uso de analizadores textuales, por lo que toda la información relativa a las normas se ha extraído manualmente.

2.2 Catálogos comerciales de cerámica

En este trabajo hemos utilizado catálogos reales de empresas cerámicas disponibles en línea o en papel. Estos catálogos incluyen descripciones con fines comerciales de las baldosas que comercializan y nos han interesado especialmente para identificar tipos de baldosas, los criterios que se han utilizado para diferenciarlas y los valores de las características que describen cada uno de estos tipos de baldosa. Este tipo de documento no resulta tan técnico como las normas de estandarización y tampoco incluye el mismo tipo de información en todos los casos. Por ejemplo, hemos descartado algunos catálogos porque apenas incluían información sobre las características técnicas de sus baldosas y se centraban únicamente en cuestiones de aspecto, de forma que el catálogo se componía de un conjunto de fotos de baldosas y una pequeña descripción sobre el tamaño y la gama de colores en los que estaba disponible. Sin embargo, los catálogos que hemos utilizado para recoger información sobre las características del producto acabado sí incluyen otro tipo de información más allá de una imagen y detalles sobre medidas y colores disponibles. Además de catálogos, hemos utilizado dos guías de una de las empresas sobre consejos y mantenimiento.

El total de documentos que hemos utilizado son los siguientes:

- ARQ 2008 (Aparici Cerámicas, 2008a)
- Catálogo Cevisama '08 (Aparici Cerámicas, 2008b)
- Catálogo general '09 (Aparici Cerámicas, 2009a)
- Catálogo novedades '09 (Aparici Cerámicas, 2009b)
- Exclusive Collection (Aparici Cerámicas, 2009c)
- Catálogo general 2009 (Halcón Cerámicas, 2009)
- Consejos de colocación (Halcón Cerámicas)
- Consejos de mantenimiento del pavimento porcelánico pulido (Halcón Cerámicas)
- Novedades 2009 (Halcón Cerámicas, 2009)
- Catálogo general de interior (Hispano Azul, 2009a)
- Crea 2009 (Hispano Azul, 2009b)
- Porcelánico Alta Gama (Hispano Azul, 2009c)
- Catálogo general Pamesa Cerámica 2008 (Pamesa Cerámica, 2008)
- Catálogo general Stylnul Cerámica 2006 (Stylnul Cerámica, 2006)
- Catálogo general Stylnul Cerámica 2007 (Stylnul Cerámica, 2007)
- Catálogo general Stylnul Cerámica 2008 (Stylnul Cerámica, 2008)
- Catálogo general Stylnul Cerámica 2009 (Stylnul Cerámica, 2009)
- Catálogo general 2008-2009 (Vives Azulejos y Gres, 2008)

Dentro de las descripciones, los catálogos incluyen descripciones de las baldosas según las normas UNE y los valores de las características más relevantes para las baldosas que comercializan.

Pavimentos gres porcelánico Porcelain tile		6,1x41,5 cm. 3.2x16.3 inches	25x45 cm. 9.8x18 inches	30x30 cm. (29.3x29.3 cm.) 11.5x11.5 inches (11.5x11.5 inches)	40x40 cm. (39.3x39.3 cm.) 15.4x15.4 inches (15.5x15.5 inches)	14,4x59,3 cm. 5.7x23.3 inches 7.6x23.3 inches	30x60 cm. (29.3x59.3 cm.) 11.5x23.3 inches (11.5x23.3 inches)	40x60 cm. (39.3x59.3 cm.) 15.7x23.6 inches (15.5x23.3 inches)	45x45 cm. (44.3x44.3 cm.) 17.4x17.4 inches (17.4x17.4 inches)	Norma de referencia Reference standard
Características técnicas / Technical Characteristics										
Proceso de fabricación Manufacture process	Manoecoción Single firing								Baldosa cerámica prensada en seco Dry pressed ceramic tile	
Grosor (mm.) Thickness	11 aprox.	B	9	9,5	10	9,5	UNE EN ISO 10545 Parte 2 / Part 2			
Longitud, anchura, espesor, rectitud de los lados, originalidad, planitud, Length, width, thickness, side straightness, orthogonality, firmness.	Cumple con los requisitos aplicables a cada modelo definidos en la norma UNE-EN 14411 para el Grupo B Ia Conforms to the requisites applicable to each model as defined in standard UNE-EN 14411 for Group B Ia								UNE EN ISO 10545 Parte 2 / Part 2	
Absorción de agua Water absorption	E < 0.5 %								UNE EN ISO 10545 Parte 3 / Part 3	
Resistencia a la flexión (N/mm.2) y fuerza de robar (N) Flexion resistance N/mm. mm. and breaking strength	> 35 > 1300								UNE EN ISO 10545 Parte 4 / Part 4	
Resistencia a la abrasión profunda. (Baldosas no esmaltadas) Resistance to deep abrasion (Un glazed tiles)	< 140mm ²								UNE EN ISO 10545 Parte 6 / Part 6	
Resistencia a la abrasión superficial (Baldosas esmaltadas) Resistance to surface abrasion (Glazed tiles)	A consultar según modelo To enquire according to de styles.								UNE EN ISO 10545 Parte 7 / Part 7	
Dilatación térmica lineal Linear thermal dilatation	< 9*10 ⁻⁴ / °C								UNE EN ISO 10545 Parte 8 / Part 8	
Resistencia al choque térmico Thermic impact resistance	Resiste la prueba It stands the test								UNE EN ISO 10545 Parte 9 / Part 9	
Resistencia al cuarteo Cocking resistance	Resiste la prueba It stands the test								UNE EN ISO 10545 Parte 11 / Part 11	
Resistencia a la helada Frost resistance	Resiste la prueba It stands the test								UNE EN ISO 10545 Parte 12 / Part 12	

Tabla 20. Extracto de tabla descriptiva del catálogo general de Vives sobre características técnicas de su pavimento de gres porcelánico (Vives Azulejos y Gres, 2008)

En la tabla 20 mostramos un fragmento del catálogo general de Vives (Vives Azulejos y Gres, 2008) donde se indican los valores de su pavimento de gres porcelánico haciendo referencia a lo que marcan las normas UNE. La extracción de este tipo de información de los catálogos nos ha permitido identificar y comparar los tipos de valor que se atribuyen a las características del producto acabado. Estos catálogos son además una fuente de ejemplos de uso reales de términos relativos a características y valores y nos han servido para enriquecer la ontología.

Los catálogos de cerámica utilizados estaban disponibles en *pdf* editable en la mayoría de los casos. Sin embargo, la mayor parte del texto estaba incrustada en forma de imágenes con lo que no se mostraban los resultados esperados al probar el análisis textual. Esto, ha supuesto que, después de una serie de intentos fallidos para procesar los documentos, la extracción de información se realizara manualmente.

2.3 El corpus TxtCeram

Además de las normas, otra información terminológica necesaria para llevar a cabo la extracción procede del corpus TxtCeram. El corpus TxtCeram es un recurso creado en la Universitat Jaume I por el Grupo Tecnolettra²⁴ dentro del proyecto homónimo TxtCeram²⁵ (Alcina, 2006; Estellés et al., 2006; Soler et al., 2006). Se trata de un corpus de lenguaje especializado del campo de la cerámica de un tamaño de 12,6 MB en formato texto (txt) y contiene 2,4 millones de palabras. Los textos que incluye son obras originales que pertenecen al campo de la cerámica.

En cuanto al tipo de información contenida en el corpus, podemos decir que contiene no solo información terminológica sino textos relevantes y explicaciones sobre los procesos y los elementos relacionados con la cerámica. De esto resulta una amplia gama de conocimiento especializado que servirá de base, entre otras, a nuestra investigación sobre clasificación de características. El corpus se compone de 34 libros (ver anexo 10) que tratan la cerámica e incluyen datos sobre: productos y tecnologías para esmalte y coloración, fabricación, usos de la baldosa, estructuras, compuestos químicos, materiales, procedimientos y procesos.

En lo que se refiere a su criterio de compilación, nuestro corpus reúne las siguientes características: es un corpus escrito, sincrónico (las obras se han publicado durante el mismo periodo, 1980-1997), especializado, textual (obras completas) y monolingüe (español).

- | | |
|-----------------------------|----------------------------------------------------|
| 1.2.3 Tipos de producto | 5. Calidad |
| 1.3. Otros productos | 5.1. Procesos de control de calidad |
| 2. Fabricación | 5.2. Maquinaria y accesorios de control de calidad |
| 2.1 Procesos | 6. Unidades de medida |
| 2.2 Maquinaria y accesorios | 7. Organismos, Instituciones |

2.4 Descripción de la baldosa cerámica de PROALSO

La Asociación Profesional de Alicatadores/Soldadores de España (PROALSO) ha desarrollado una serie de manuales de formación sobre los que fundamenta los contenidos para la obtención del Carnet Profesional de Alicatador Soldador. El *Manual de Actualización en materiales y técnicas de colocación de recubrimientos cerámicos* (PROALSO, 2011) contiene un capítulo dedicado a la clasificación y selección de la baldosa cerámica, que incluye una tabla resumen (tabla 21). En este capítulo, los autores describen distintos tipos de baldosa relacionándolos con la clasificación de la norma UNE-EN 14411 (AENOR, 2007). La información contenida en este capítulo se ha utilizado en este trabajo como ejemplo para la extracción de tipos de baldosa desde el plano comercial y para la consulta sobre descripción de baldosas y características.

²⁴ <http://tecnolettra.uji.es/es/>

²⁵ TXTCeram: Extracción semiautomática y análisis conceptual formal de términos de la cerámica a partir de un corpus electrónico. Su eficacia y utilidad en la mediación lingüística. Generalitat Valenciana (GV05/260)



Cermet Profesional Alicatador Solador, Aplicaciones Convencionales
Proal: Asociación Profesional de Alicatadores/Tilostadores

CLASIFICACIÓN COMERCIAL DE LAS BALDOSAS CERÁMICAS			
	Denominación Comercial	Segun Norma UNE-EN 14411	Definición
Barro cocido	Baldosas tierra cocida	AIIb UGL 6%<E ≤ 10%) AIII, UGL E>10%	Baldosas cerámicas de soporte coloreado y poroso sin aplicación de vidriados ni decoraciones. Conformación por moldeo manual, extrusión y prensado. Tratamientos superficiales sobre producto acabado para embellecimiento y protección a las manchas. Mayoritaria utilización en pavimentos. El gran éxito comercial ha propiciado la espectacular ampliación de repertorio: en formatos, aplicación de vidriados, insertos e incrustaciones, combinación con piedra natural, etc.
	Cotto Pavimento rústico Pavimento terracotta Toba Toba rustica		
Baldosin catalán	Baldosin rojo mate Rajola Rasilla	AIIb UGL 6%<E ≤ 10%) AIII, UGL E>10%	Baldosas cerámicas de soporte coloreado y poroso sin aplicación de vidriados ni decoraciones. Conformación por extrusión. Se utiliza para solado de terrazas, balcones y porches, con frecuencia en combinación con alambillas.
Azulejo	Azulejo de pasta blanca Rajola Baldosas de mayólica Baldosas de loza Revest. de bicocción Revest. monococción Maniseta Taufell	BIII, GL, E>10%	Baldosas cerámicas de soporte coloreado, blanco o blanco/grisáceo y poroso, con aplicación de vidriados y decoraciones. Tratamientos decorativos a baja temperatura. Conformación por prensado en semiseco y sometido a una o varias cocciones (monococción, bicocción, bicocción rápida, tercer fuego, etc.) Mayoritaria utilización en revestimientos interiores. Puede utilizarse en exteriores no sometidos a riesgo de helada. Gran variedad de productos por color del soporte (brazochis), proceso de fabricación, formatos y tratamientos superficiales cerámicos con fines decorativos. Se complementan con piezas especiales.

Tabla 21. Extracto de la tabla de clasificación comercial de las baldosas cerámicas de la asociación PROALSO (PROALSO, 2011)

2.5 El Diccionario Cerámico Científico-práctico

Como recurso para obtener información y para documentarnos sobre el significado de algunos conceptos hemos utilizado el Diccionario Cerámico Científico-práctico (Guillem Monzonís y Guillem Villar, 1987). Este diccionario, originalmente en papel, se digitalizó dentro del proyecto CREC²⁶ del Grupo Tecnolettra y se ha utilizado con dos propósitos y en dos formatos: por un lado en formato *txt* como parte del corpus TxtCeram y por otro en formato Word para su consulta y acceso.

El diccionario es una obra realizada por dos expertos del campo de la cerámica y constituye un material de referencia muy valioso. 2396 términos, algunos de los cuales son características, con su definición y sus equivalentes a inglés, francés y alemán, han servido de apoyo en este proyecto.

²⁶ Proyecto de mejora educativa «CREC: Creación de recursos lingüísticos electrónicos», financiado por la *Unitat de Suport Educatiu* (USE) de la Universitat Jaume I, 2004-05, y proyecto de mejora educativa «CREC II: Creación de recursos lingüísticos electrónicos», financiado por la *Unitat de Suport Educatiu* (USE) de la Universitat Jaume I, 2005-06.

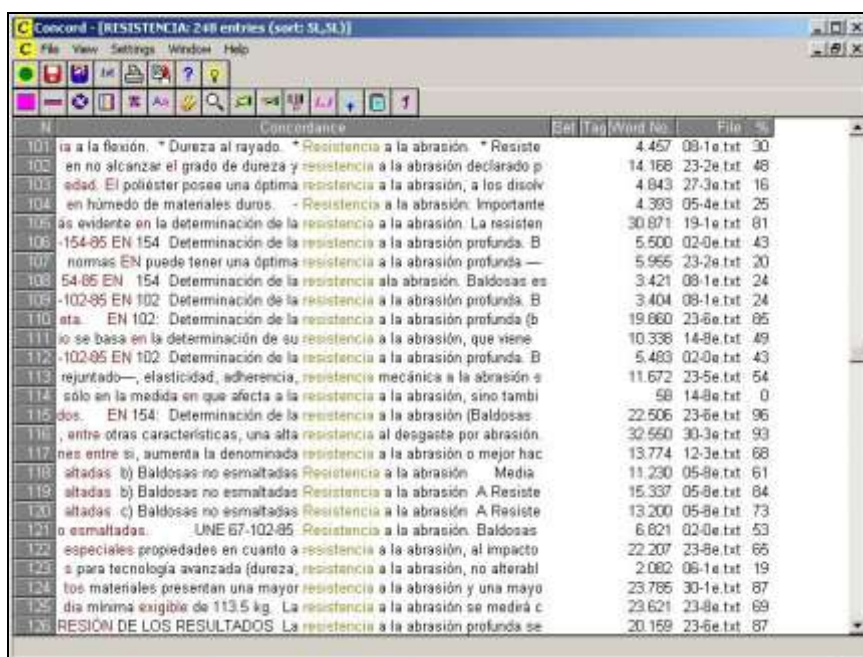
3 Herramientas informáticas

3.1 WordSmith

WordSmith es una aplicación popular a utilizada en el análisis textual que ha sido diseñada por Mike Scott. En nuestro trabajo hemos utilizado la versión 3.0 de este sistema que integra tres herramientas: Wordist, Keywords y Concord.

WordList es una herramienta que, a partir de un corpus, proporciona listas de palabras o grupos de palabras de un texto en orden alfabético o en orden de frecuencias. KeyWords se utiliza para extraer palabras clave a partir de la comparación estadística de dos corpus de especialidades distintas, mediante la contrastación de segmentos o palabras que no tienen en común. Y finalmente, Concord nos permite buscar palabras en contexto.

Esta última herramienta, que es la que hemos utilizado en nuestro trabajo, permite la creación de listados de términos en contexto a partir de la búsqueda de una o más palabras, es decir, permite la búsqueda de unidades complejas (figura 70). La herramienta proporciona información acerca del número de ocurrencias encontradas en el corpus, su contexto, el porcentaje de aparición en el corpus, etc. Además, una vez realizada la búsqueda es posible buscar patrones de orden mediante un sistema basado en el cálculo de probabilidades a partir del corpus.



The screenshot shows the Concord software interface with a search window titled 'Concord - [RESISTENCIA: 246 entries (sort: SL,SL)]'. The main window displays a list of search results with columns for 'N', 'Concordance', 'Est', 'Tag', 'Word No.', 'File', and '%'. The results show various occurrences of the terms 'resistencia' and 'abrasión' in different contexts, such as 'Resistencia a la abrasión' and 'Resistencia mecánica a la abrasión'.

N	Concordance	Est	Tag	Word No.	File	%
101	ia a la flexión. * Dureza al rayado. * Resistencia a la abrasión. * Resiste			4.457	09-1e.txt	30
102	en no alcanzar el grado de dureza y resistencia a la abrasión declarado p			14.168	23-2e.txt	48
103	edad. El poliéster posee una óptima resistencia a la abrasión; a los disolv			4.843	27-3e.txt	16
104	en húmedo de materiales duros. - Resistencia a la abrasión. Importante			4.383	05-4e.txt	25
105	as evidente en la determinación de la resistencia a la abrasión. La resisten			30.871	19-1e.txt	81
106	-154-85 EN 154 Determinación de la resistencia a la abrasión profunda. B			5.500	02-0e.txt	43
107	normas EN puede tener una óptima resistencia a la abrasión profunda —			5.955	23-2e.txt	20
108	54-85 EN 154 Determinación de la resistencia ala abrasión. Baldosas es			3.421	08-1e.txt	24
109	-102-85 EN 102 Determinación de la resistencia a la abrasión profunda. B			3.404	08-1e.txt	24
110	ata. EN 102: Determinación de la resistencia a la abrasión profunda (b			19.660	23-6e.txt	95
111	o se basa en la determinación de su resistencia a la abrasión, que viene			10.338	14-8e.txt	49
112	-102-85 EN 102 Determinación de la resistencia a la abrasión profunda. B			5.483	02-0e.txt	43
113	rejuntao—, elasticidad, adherencia, resistencia mecánica a la abrasión s			11.672	23-5e.txt	54
114	sólo en la medida en que afecta a la resistencia a la abrasión, sino tambi			59	14-8e.txt	0
115	dos. EN 154: Determinación de la resistencia a la abrasión (Baldosas			22.506	23-6e.txt	96
116	, entre otras características, una alta resistencia al desgaste por abrasión			32.550	30-3e.txt	93
117	nes entre si, aumenta la denominada resistencia a la abrasión o mejor hac			13.774	12-3e.txt	68
118	altadas b) Baldosas no esmaltadas Resistencia a la abrasión Media			11.230	05-8e.txt	61
119	altadas b) Baldosas no esmaltadas Resistencia a la abrasión A Resiste			15.337	05-8e.txt	84
120	altadas c) Baldosas no esmaltadas Resistencia a la abrasión A Resiste			13.200	05-8e.txt	73
121	o esmaltadas. UNE 67-102-85 Resistencia a la abrasión. Baldosas			6.821	02-0e.txt	53
122	especiales propiedades en cuanto a resistencia a la abrasión, al impacto			22.207	23-6e.txt	65
123	s para tecnología avanzada (dureza, resistencia a la abrasión, no alterabl			2.082	06-1e.txt	19
124	los materiales presentan una mayor resistencia a la abrasión y una mayo			23.785	30-1e.txt	87
125	dia mínima exigible de 113,5 kg. La resistencia a la abrasión se medirá c			23.621	23-8e.txt	69
126	RESIÓN DE LOS RESULTADOS La resistencia a la abrasión profunda se			20.159	23-6e.txt	87

Figura 70. Búsqueda de “resistencia” y “abrasión”

En nuestro trabajo, hemos utilizado las ventajas de esta última herramienta y la hemos aplicado utilizando de fuente el corpus TxtCeram. La herramienta ha servido para extraer el número de ocurrencias de cada una de las características de la baldosa y para extraer contextos definitorios e información conceptual que pueda ayudar a su clasificación.

3.2 Protégé 4 y HerMiT

Protégé (véase §II.10.1) es la herramienta que hemos utilizado para elaborar la ontología de características de la baldosa cerámica y HerMiT (véase §II.10.1.12) el razonador de ontologías OWL 2 que hemos utilizado para verificar la coherencia y consistencia de los datos representados.

El editor de ontologías Protégé se ha utilizado para representar en forma de clase, propiedad, instancia o expresión toda la información que recoge nuestro marco conceptual y la información obtenida en la etapa de extracción. Como hemos visto anteriormente, este editor nos permite representar los conceptos, utilizando las distintas ventanas y menús, a través de clases, propiedades o expresiones. Hemos utilizado una gran mayoría de las opciones que permite el editor, principalmente para realizar las siguientes tareas:

- Creación de clases y estructura de clases
- Creación de instancias de clases
- Creación de propiedades, asignación de rangos y dominios
- Creación de anotaciones sobre entidades OWL 2
- Descripción de clases a través de expresiones
- Búsqueda de conceptos en la ontología para validación y verificación
- Verificación de la consistencia de la ontología
- Consultas en DL Query sobre propiedades y clases

Para introducir las expresiones que describirán las clases de la ontología hemos utilizado la *Manchester OWL 2 Syntax* (véase II.8.2).

Una herramienta muy utilizada que viene en la aplicación de Protégé 4 ha sido *Create Class Hierarchy*. Esta herramienta nos permite copiar una estructura de conceptos tabulada (elaborada en un procesador de textos corriente) y pegarla en la ventana de Protégé. Esta estructura se interpretará como estructura de clases y subclases y se incorporará a la ontología (figura 71). Esto evita, por ejemplo, tener que crear cada clase manualmente, una por una, y asignar la superclase correspondiente.

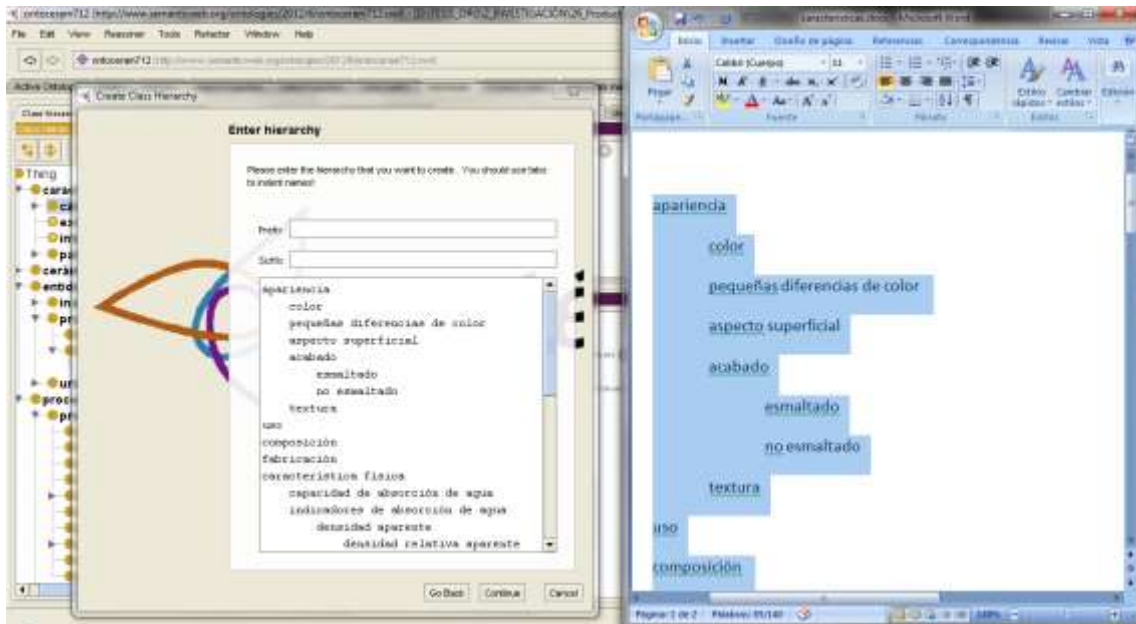


Figura 71. Creación de clases y subclases desde *Create Class Hierarchy* en Protégé y ejemplo de estructura desde Word

Para visualizar la información formalizada sobre los conceptos en la ontología en forma de grafo y representarla en este documento hemos utilizado las herramientas integradas en Protégé *OWL 2 Viz* y *Ontograf* (véase II.10.1.1). Estas herramientas nos permiten visualizar las clases y las relaciones entre clases a través de grafos basándose en la jerarquía representada o inferida según se desee en el caso de *OWL 2 Viz*. *Ontograf* además de permitir la representación de la jerarquía de clases, nos permite representar la relación entre una clase y otra basándose en otro tipo de propiedad. La utilización de una herramienta u otra para representar los grafos en este documento ha dependido del aspecto que quisiéramos reflejar y de criterios estéticos relativos a una mejor visualización del contenido del grafo (figura 72).

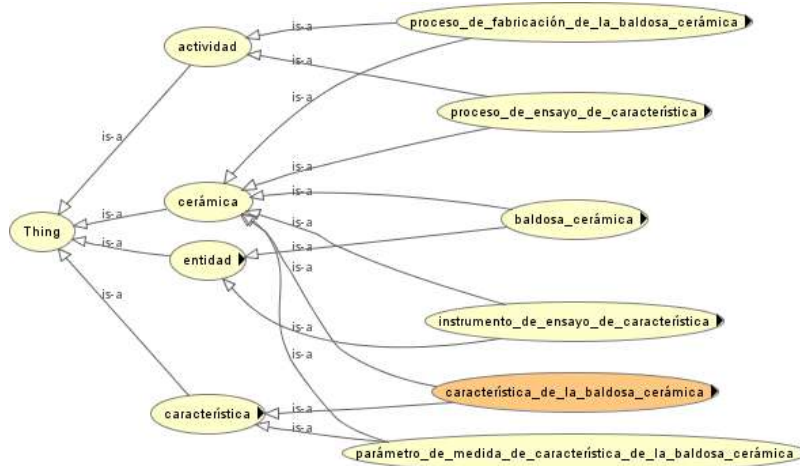


Figura 72. Ejemplo de grafo generado con *OWL 2 Viz*

Además, hemos utilizado el complemento *Matrix Views* para facilitar el trabajo de edición e introducción de descripciones sobre las clases de la ontología. *Matrix Views* es un plug-in desarrollado por Nick Drummond y Timothy Redmond dentro del proyecto CO-ODE de la Universidad de Manchester (Drummond et al., 2010). Este complemento añade hasta tres

pestañas para facilitar la edición en Protégé. Cada pestaña nos permite añadir descripciones sobre un tipo de elemento de la ontología (*class*, *property* o *individual*) en forma de matriz. La ventaja de este complemento con respecto a elaborar la descripción utilizando las pestañas por defecto de Protégé recae en que todos los datos según tipos se organizan en una misma matriz de modo que podemos ver, por ejemplo, todas las propiedades disponibles para describir una clase y realizar la descripción desde la misma ventana. Esto evita tener que abrir, editar y cerrar una ventana para cada propiedad que queramos describir de una clase, tal y como se realizaría si usáramos la metodología tradicional de Protégé (§II.10.1.5), en su lugar podemos rellenar toda la información desde la misma tabla. Además, cuando asignamos una descripción a una clase, podemos observar cómo se hereda en las subclases y verificar si el contenido es correcto. En la figura 73 se muestra un extracto de la descripción del tipo de característica (intrínseca/extrínseca) en las características de la ontología. Los campos que no están entre paréntesis se han introducido manualmente, los que aparecen entre paréntesis muestra que son valores heredados por la clase.

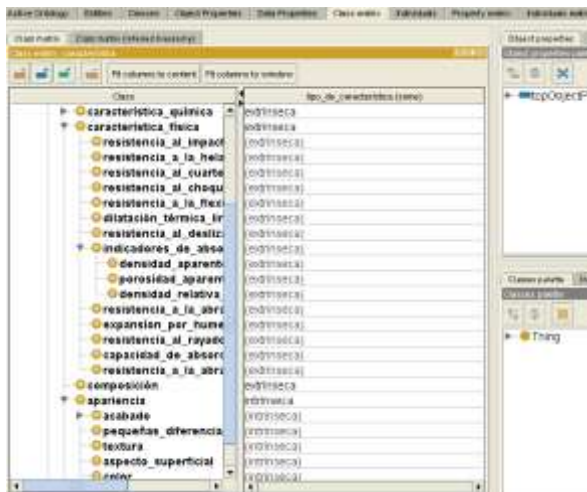


Figura 73. Ejemplo de *class matrix* con la propiedad 'tipo_de_característica' y tipo de descripción 'some'.

El razonador HerMiT (véase §II.10.1.12), integrado en el editor Protégé, se ha utilizado a lo largo de la fase de modelado y de introducción de datos para verificar que no existían errores de coherencia. El razonador se puede activar y desactivar en cualquier momento desde el menú *Reasoner* (figura 74) a través de *Start reasoner* (si lo vamos a utilizar por primera vez en esa sesión de trabajo) o *Synchronize reasoner* (si queremos que se sincronice con la nueva información que hayamos implementado tras la última aplicación del razonador).



Figura 74. Menú de uso del razonador en Protégé 4

Es importante remarcar que el razonador en ningún momento modifica la estructura original de clases representada en la ontología sino que muestra una estructura alternativa inferida. Este conocimiento implícito puede visualizarse en diferentes puntos de la interfaz, por ejemplo, en la sub-ventana paralela *Class hierarchy (inferred)* o en el apartado de *Inherited*

anonymous classes de la sub-ventana de descripción (figura 75). Al poder comparar la clasificación original de conceptos (realizada por nosotras) con la clasificación inferida por el razonador a partir de la información representada, detectamos qué información ha sido inferida y no será necesario explicitar manualmente, lo que permite ahorrar el tiempo de introducción de datos.

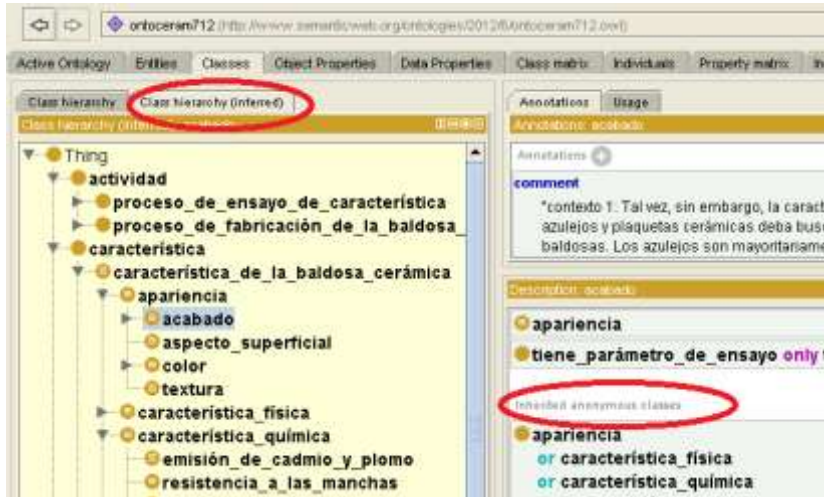


Figura 75. Visualización de información inferida por HermiT.

El uso de HermiT en la fase de construcción de la ontología también nos permite comprobar si las descripciones que estamos realizando son las esperadas o estamos pasando por alto algún detalle que genera inconsistencias en la ontología. Por ejemplo, si hemos asignado una misma subclase a clases descritas como disjuntas (que no pueden compartir miembros), el razonador nos indica que existe un error en ese punto. Igualmente, si hemos introducido una expresión que entra en conflicto con otra o si se infiere información no deseada, el uso del razonador nos permite detectar este tipo de casos y corregirlos.

Por ejemplo, en la ontología se ha descrito que el valor de 'tiene ortogonalidad' es un número de tipo decimal (*double*). Si en la clase de baldosa cerámica 'grupo A' declaramos que la clase tiene como valor de ortogonalidad un número entero (1 en lugar de 1.0), el razonador nos indicará que existe un error (figura 76).

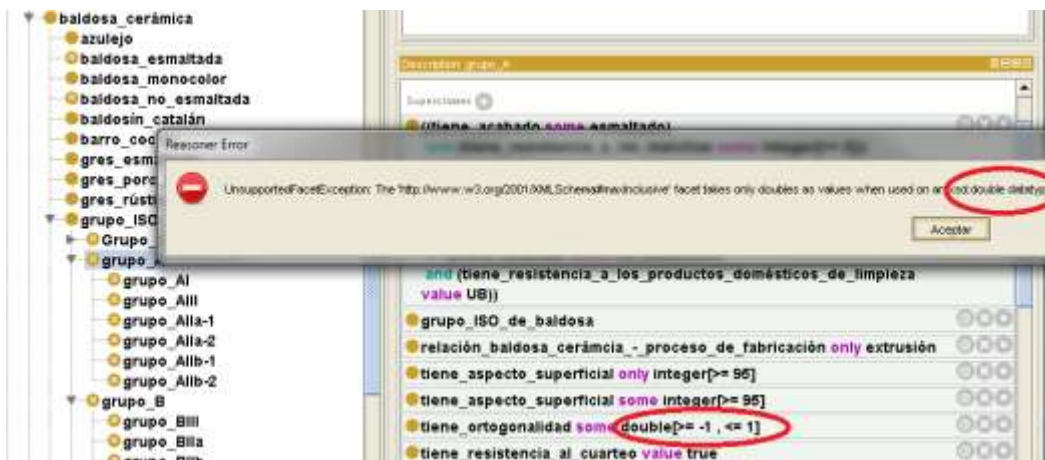


Figura 76. Ejemplo de error por valor mal expresado

Finalmente, el razonador HerMiT también se ha utilizado en la fase final del trabajo para hacer pruebas de consultas sobre la ontología. Por ejemplo, para buscar qué baldosas se describen a través de la característica ‘capacidad de absorción de agua’. En este caso, el razonador HerMiT se activa al realizar la consulta en la pestaña de consultas de Protégé DL Query (véase §II.10.1.11).

4 Identificación, descripción y estructuración de los elementos que ayudan a describir una característica

4.1 Introducción

Este apartado describe el trabajo realizado en la fase de identificación, descripción y estructuración de las características de la baldosa cerámica y de los conceptos necesarios para describirlas. Nos centramos en identificar qué elementos aparecen en la descripción y definición de una característica en la bibliografía. Esta fase se ha llevado a cabo de forma simultánea a la fase de extracción de las características de la baldosa y su información conceptual. La información teórica sobre la naturaleza de cada elemento se ha contrastado con casos prácticos encontrados durante el proceso de extracción. De este modo, se ha contribuido a agregar más detalles sobre la naturaleza de los conceptos que componen la ontología y al mismo tiempo se ha garantizado la viabilidad de este trabajo.

Los enfoques que hemos analizado se pueden consultar a lo largo de la sección de fundamentos teóricos de este trabajo (§II). Hemos aislado cada uno de los criterios que se ha utilizado para clasificar o describir características o conceptos en general en la bibliografía y que considerábamos que podían contribuir a la creación de nuestra ontología. Primero, hemos ido anotando cada elemento que podría servir para describir las características, incluyendo información sobre otros conceptos relacionados, el autor o escuela que lo trataba, etc. El siguiente paso ha consistido en comparar los elementos encontrados para comprobar si trataban de aspectos diferentes o compartían un mismo criterio. Por ejemplo, los distintos enfoques que tratan la distinción entre nombre de una característica y valor de una característica (Meyer, 1997:100; Kageura, 1997:77; Thomsen, 1998:356) que hemos analizado en el apartado §II.2.2.

Estos criterios se han puesto en común con las necesidades y objetivos de nuestro proyecto y con la información detectada durante la fase de extracción. Esta puesta en común da lugar al siguiente conjunto de elementos considerados necesarios para llevar a cabo la descripción de las características y su posterior formalización en la ontología.

La representación de los elementos que forman parte de la descripción de una característica y de las relaciones que se establecen entre estos elementos tiene como objetivo la propia descripción de características. Es decir, nos hemos planteado su representación con fines descriptivos y más concretamente, descriptivos de la propia característica. Al representar formalmente todos los elementos que integran una característica, junto con sus relaciones, estamos formalizando una característica y cumplimos así con uno de los objetivos de nuestro trabajo.

Al mismo tiempo, esta descripción de características de manera detallada y organizada nos ha permitido identificar cómo interactúan los diferentes elementos que participan en la representación de una característica de modo que podemos establecer una serie de pautas o requisitos a considerar cuando representamos características. Estas pautas o requisitos hacen referencia a la información que necesitamos saber de una característica y también a los distintos tipos de información que podemos manejar cuando trabajamos en la descripción de una característica.

Otro aspecto significativo de este trabajo es que nos hemos centrado en aprovechar las ventajas que supone el uso de un sistema de ontologías para la representación de características. Por eso, consideramos todos aquellos aspectos o información relativa a una característica susceptible de representarse en una ontología de modo que podamos aprovechar al máximo las funcionalidades de la propia ontología (principalmente, la inferencia a través de razonadores). Pero antes de llegar a este punto, vamos a describir los elementos que utilizaremos para describir una característica.

Orientado desde las características, el esquema de descripción se compone de los siguientes elementos:

- Nombre de la característica
- Valor
 - Tipo de valor
- Ámbito
- Tipología
 - Intrínseca o extrínseca
- Relaciones
 - Esencial, no esencial o distintiva para la baldosa cerámica
 - Hipónimos- Hiperónimos
 - Relación de meronimia
 - Relación característica – proceso de ensayo
 - Sinonimia y equivalencia en otra lengua
- Características de las características
 - Característica equivalente
 - Instrumento de ensayo de la característica
 - Parámetro de ensayo de la característica
 - Unidad de medida
- Explicación en lenguaje natural
 - Contextos
 - Fuente de la característica
 - Observaciones

Estos elementos activan la presencia de otro tipo de conceptos además de las características, como son entidades, relaciones o actividades.

A continuación, explicamos qué entendemos que son cada uno de ellos, qué papel tendrán en la formalización de una característica, qué relaciones mantienen con otros elementos y qué consideraciones debemos tener en cuenta de cara a su representación en la ontología.

4.2 Característica

En este trabajo, una **característica** se observa como un concepto que sirve para distinguir otros conceptos; es decir, que denota las facetas que tienen otros conceptos que bien pueden ser entidades, relaciones, otras características o actividades. Entendemos ‘característica’ como el tipo de concepto que ayuda a describir un concepto y que denota las diferencias o semejanzas entre conceptos. Las características tienen un doble rol en un sistema de conceptos; por una parte, constituyen conceptos en sí mismos que hacen referencia a los rasgos que describen otros conceptos; por otra, constituyen un elemento descriptor que se utiliza para expresar los rasgos de otros conceptos. Así, podemos estudiar la característica ‘color’ como concepto en sí mismo y describir sus características, su ámbito de uso o objetos que puede describir, y podemos acercarnos a la característica ‘color’ como un elemento que permite atribuir esa característica a un concepto y relacionarlo con un valor.

En el plano conceptual, una característica se distingue del resto de conceptos porque persigue la descripción de un rasgo de un concepto. No obstante, en el plano de representación de conceptos, donde la característica se observa en su rol descriptor como un binomio atributo-valor, la representación de una relación y una característica sólo se diferencia por la etiqueta que describe el atributo. Por eso, podemos afirmar que de la expresión ‘el libro es de color rojo’ se deduce que ‘color’ es una característica que está describiendo el concepto ‘libro’ con el valor ‘rojo’, y al mismo tiempo aceptamos que existe una relación de tipo objeto-característica entre el concepto ‘libro’ y el concepto ‘rojo’.

En nuestro objetivo por representar las características de manera formal, partimos de una distinción inicial entre el nombre que recibe una característica y los valores concretos que puede tener.

4.2.1 Nombre de la característica

El **nombre de la característica** es la denominación que recibe una faceta que caracteriza un concepto. Por ejemplo, el ‘color’ de una baldosa se refiere a una faceta presente en el concepto ‘baldosa’ relativa a una impresión debida por el reflejo de la luz. En relación con los valores de una característica, el nombre de una característica funciona como una categoría que sirve para agrupar e identificar al conjunto de variantes posibles para una misma característica. Por ejemplo, ‘color’ sirve para agrupar los valores ‘rojo’, ‘verde’, etc. Por esta razón, a veces el nombre de la característica se observa como una representación abstracta de todos los posibles valores de la característica, es decir, cuando aludimos al ‘color’ de una baldosa, estamos aludiendo a los posibles valores de la baldosa. Y por eso, **cuando queremos hacer referencia a una característica de forma abstracta hacemos uso del nombre de ésta**. Esto nos lleva a determinar que el nombre de la característica es la forma en la que representamos el concepto abstracto que representa una característica.

Al mismo tiempo, el nombre de una característica limita el significado del concepto representado por un valor, por ejemplo, cuando ‘rojo’ está vinculado a ‘color’, ‘rojo’ representa la faceta del tono impreso por la luz, en cambio, cuando ‘rojo’ acompaña a ‘ideología’ el concepto que se representa es distinto.

Mediante el nombre de la característica representamos la información relativa a una característica en concreto y, como veremos más adelante, el nombre es un elemento que nos permite establecer relaciones entre ésta y otros conceptos de una ontología y mediante el cual

describimos las propiedades de la característica (si es intrínseca o extrínseca, si es esencial o distintiva, si tiene antónimos, etc.). Además, cuando la característica se observa como **elemento descriptor**, el nombre de la característica **acompañará a un valor o valores** en la ontología que harán referencia a una faceta concreta de un determinado concepto. Al representar toda esta información en nuestra ontología, aportamos información sobre qué:

- denominación o denominaciones puede recibir una característica porque la haremos explícita en nuestra ontología
- conceptos pueden ser descritos por las características porque estableceremos relaciones entre éstas y los conceptos que describen
- tipo de características están siendo utilizadas para describir otros conceptos porque podremos observar qué relaciones se activan o desactivan dependiendo de qué conceptos y podremos añadir información sobre la característica a través del elemento ‘nombre de característica’.

El nombre de la característica tiene una **relación con los valores** a los que agrupa que podemos entender como una relación **de tipo lógico**, la relación vincula el nombre de característica y sus posibles ejemplares: los valores. Un ejemplar de ‘color’, por ejemplo, es ‘rojo’. Existe también una relación de hiponimia entre los nombres de característica más generales y aquellos más específicos. Por ejemplo, ‘apariencia’ es más general que ‘color’ y entre ellos mantienen una relación de tipo *es-un* (hiponimia). El nombre de la característica, como representante denominativo del concepto abstracto de la característica, mantiene **relaciones con las denominaciones de los conceptos a los que describe** y éstas se pueden entender como relaciones esenciales o relaciones distintivas (describimos estas relaciones más adelante). Además de las relaciones, el nombre de la característica, como representante del concepto abstracto de la característica, será de tipo **extrínseco o intrínseco** (describimos este aspecto más adelante), y este aspecto se le atribuirá también al concepto.

El nombre de la característica podrá ser sinónimo y antónimo de otros nombres de característica que hacen referencia al mismo concepto abstracto o a su opuesto, respectivamente. Por eso, mediante el nombre de la característica se pueden establecer también relaciones de sinonimia y antonimia, como veremos más adelante.

Asimismo, el nombre de la característica está vinculado a una **lengua** y existirán nombres de característica que hagan referencia al mismo concepto en otras lenguas.

4.2.2 Valor de la característica

El **valor de la característica** es un concepto mediante el cual se representa un **matiz concreto de la característica**, estos conceptos se pueden observar, desde el punto de vista representacional, como datos que se agrupan bajo categorías que los identifican y se interpretan como gradaciones, ejemplares o variantes de una característica.

Desde la ingeniería del conocimiento, observamos muestras de que el valor de una propiedad **se puede tratar como un tipo de dato cuya representación depende de su naturaleza**. Desde este punto de vista, los valores de una característica podrán ser de distintos tipos (un valor numérico, un valor booleano, etc.) y se da prioridad a clasificarlos según una tipología u otra con fines de facilitar su representación en un sistema informático y su procesamiento.

La identificación y representación de los valores de una característica resulta fundamental para describir la naturaleza de una característica y poder hacernos una idea de su aplicación. Los valores constituyen ejemplos concretos relativos a la característica y mediante ellos identificamos la extensión de la característica.

Además, cuando describimos un concepto, es el valor de la característica el que aportará la información que nos permita diferenciarlo de otros. Con lo que su representación de manera estructurada y estableciendo vínculos con el nombre de la característica y los conceptos donde está presente, nos permitirá visualizar qué conexiones tiene una característica y de qué manera influye en uno u otro concepto.

El hecho de clasificarlos en una tipología siguiendo las pautas de la Ingeniería del conocimiento, nos permite además aportar información sobre la naturaleza del valor y de la misma característica. Por ejemplo, si formalizamos que una característica como ‘altura’ tiene un valor que se compone de un número y una unidad de medida que se expresa en centímetros, estamos expresando que esta característica tiene una relación con aquellos conceptos que se puedan medir utilizando la unidad de medida ‘centímetros’; o ir más allá e inferir que esta característica podría estar presente en todas las entidades físicas que se puedan medir utilizando esta unidad de medida.

Siguiendo con la tipología de valores, debemos remarcar que existen valores simples, que aluden a un sólo concepto (por ejemplo, ‘rojo’) y valores complejos, que se pueden entender como términos complejos ya que se constituyen de más de un concepto, por ejemplo, tres centímetros, que alude un concepto ‘número’ y otro ‘unidad de medida’ o tres milímetros por decilitro cuadrado. En este último caso, debemos decidir si trataremos ‘milímetro por decilitro cuadrado’ como un todo o por separado. Esta y otras situaciones donde según el caso concreto resulta interesante expresar una estructura más o menos compleja para describir un valor, nos han llevado a recoger el máximo de información sobre cada uno de los valores que queremos representar en nuestra ontología, con el fin de identificar un conjunto de tipos de valores y de relaciones que activan o se requieren para ayudar a describir una característica.

Con esto, queremos decir que la tipología que ha resultado se inspira en los trabajos de ingeniería del conocimiento y más concretamente en el lenguaje de representación de ontologías OWL 2, pero se fundamenta principalmente en los grupos de valores que hemos ido identificando durante la fase de extracción y se ha elaborado *ad hoc* para este proyecto.

Hacemos una primera distinción general entre aquellos valores que se representan mediante palabras o términos de aquellos que se representan mediante (o contienen) números.

Dentro del primer grupo distinguimos un subgrupo de valores que son términos o palabras. Estos valores se representarán en forma de clase o instancia en nuestra ontología, por ejemplo, los tipos de color o los tipos de forma. En este subgrupo también se incluyen **abreviaciones, siglas o símbolos** basados en la abreviación de palabras o términos. Por ejemplo, valores de ‘resistencia a los ácidos en concentraciones débiles’ son ULA (sin efecto visible), ULB (efectos visibles en lados cortados), ULC (efectos visibles en lados cortados y no cortados, y en la cara vista); estos valores representan las iniciales de baldosa no esmaltada (Unglazed) con resistencia a los ácidos en concentraciones débiles (**Low**) tipo **A,B** o **C** (siendo C el nivel más bajo dentro de la escala de A a C).

El segundo grupo que hemos distinguido es el de los valores **booleanos**. Existen casos donde la característica interesa porque está presente o ausente, se tiene o no se tiene, y no importa el valor concreto que pudiera tomar. Por ejemplo, cuando hablamos de la ‘resistencia a la helada’ de una baldosa, desde el campo de la cerámica, no se alude a cuántos grados bajo cero resiste sino simplemente, a si cumple las condiciones necesarias para tener dicha característica o no y, por eso, en los catálogos o en los manuales encontraremos como valores de esta característica: ‘resiste/no resiste’ o ‘sí/no’. Para este tipo de valores, resulta interesante saber que se trata de valores Booleanos pero también se trata de representar qué expresiones o términos se utilizan para hacer referencia al valor Booleano (si suele decirse ‘presente’ o ‘resiste’ o ‘sí’, etc.) ya que así aportamos información lingüística que puede ser interesante para el terminólogo.

Un tercer tipo de valores que podemos encontrar son los valores numéricos, donde hemos identificado dos grandes tipos:

- **Valores numéricos seguidos de unidades de medida.** Este tipo de valor se ha identificado durante la fase de extracción y resulta relevante desde el punto de vista de su formalización ya que vincula la característica a otro tipo de concepto, la unidad de medida que corresponda. Además, como hemos visto anteriormente, puede darse el caso de que el valor incluya una combinación de unidades de medida, lo que aumenta la complejidad de su formalización.
- **Valores numéricos porcentuales.** Existen características cuyo valor expresa el porcentaje en que se podría determinar la presencia de dicha característica en un concepto. Estos porcentajes surgen de procesos o métodos de medición concretos y por tanto, puede ser interesante identificar el proceso o método utilizado para determinar dichos valores. Por ejemplo, la ‘dilatación por humedad’ de una baldosa se expresa en valores porcentuales y el valor indica el porcentaje de ‘aumento de volumen de un material causado por la absorción de agua debida a la humedad’. Esta información, aparecerá en la descripción de la característica en forma de información contextual.

Independientemente del tipo de valor, encontramos casos donde es necesario aplicar una **restricción de valor máximo/mínimo** o de **lista cerrada** dentro de los límites posibles para un determinado concepto. Por ejemplo, cuando hablamos de la ‘resistencia al rayado’ de una baldosa utilizamos una escala denominada ‘Escala de Mohs’ que atribuye un valor del 0 al 10 o en el caso de las características cuyos valores son siglas o símbolos constituidos a partir de unidades léxicas siempre encontramos que se trata de listas cerradas de valores posibles. En estos casos, podemos afirmar que los valores están limitados. Sin embargo, cuando hablamos de valores para la característica ‘color’, los valores que se agrupen pueden aumentar o disminuir a medida que vamos describiendo nuevos objetos o dependiendo del nivel de concreción del color.

Esta propiedad de **valor limitado o ilimitado** puede aplicarse **por la naturaleza** de la característica, como en el caso de los valores según la Escala de Mohs o por necesidades de la ontología, por ejemplo, podemos añadir restricciones sobre el porcentaje de la capacidad de absorción y utilizar estas **restricciones para distinguir entre tipos de baldosa**.

Otro tipo de **restricción** de los valores aparece cuando necesitamos limitar el **número** de valores posibles que se pueden **enlazar** mediante un mismo nombre de característica a una misma entidad. La descripción de entidades que representan objetos reales requiere que, por

ejemplo, necesitemos declarar que una baldosa no pueden tener brillo y a la vez no tener brillo o tener una resistencia al rayado de 2 Mohs y a la vez tenerla de 8 Mohs. Por eso, resulta necesario señalar que existen características que cuando describen entidades concretas, no pueden tomar más de un valor. Sin embargo, también existen características que cuando describen a una entidad, sí pueden tomar más de un valor, como es el caso de la característica ‘color’ para baldosa, una misma baldosa puede tener más de un color.

En relación con el grado de especificación de los valores de una característica, cuando estudiábamos la representación en la ontología de los valores de tipo numérico nos hemos cuestionado la necesidad de recoger datos concretos sobre valores de características o incluir solamente información sobre el tipo de valor. Por ejemplo, decidir si conviene representar que el revestimiento de tipo Grupo BIII tiene una resistencia a las manchas con valor ‘3’ o si desde el punto de vista del terminólogo bastaba con representar que esa característica (resistencia a las manchas) tiene un valor de tipo numérico.

La conclusión de este debate ha sido que nuestra ontología debe recoger los valores concretos que pueden tomar las clases de baldosa cerámica. Esta decisión está motivada por razones de descripción, ya que para distinguir entre tipos de baldosa, es necesario conocer los valores que toman algunas características, al ser estos valores el aspecto distintivo entre una clase y otra.

4.2.3 **Propiedades matemáticas de las características**

En Ingeniería del Conocimiento, toman de las matemáticas conceptos que aluden a la simetría o a la transitividad de las propiedades que se utilizan para describir conceptos. Estos conceptos también están presentes en Terminología cuando se habla de las relaciones conceptuales. En el campo de las características, cuando se representan como descriptoras de un concepto, solemos concebirlas como un nombre de característica que se atribuye al concepto descrito y que va acompañado por el valor o valores de la característica para ese concepto. Desde esta posición que acerca la representación de una característica describiendo a un concepto a la utilizada para expresar una relación entre conceptos, podemos afirmar que en una característica también incurren algunas de las propiedades que describe la matemática y que ya se han aplicado a las relaciones conceptuales anteriormente en Terminología.

Una característica es **funcional** cuando ésta sólo puede aparecer tomando un valor una vez en la descripción de un concepto. Es decir, si indicamos que ‘color’ tiene la propiedad de ser funcional, en la descripción de una baldosa ‘color’ sólo podrá aparecer una vez con un determinado color como valor de ésta. En cambio si una característica no es funcional puede aparecer en la descripción de un concepto con más de un valor y se le atribuirán distintos valores. En el caso de ‘color’ estaríamos diciendo que una baldosa puede tener más de un color. Esta propiedad resulta útil para describir una característica ya que permite indicar si puede tener más de un valor o no para un mismo concepto. Por ejemplo, en el caso de características como ‘resistencia a las manchas’ cuyo valor debe ser uno y sólo uno entre una escala del 1 al 5, al definir que es funcional estamos indicando precisamente esa información, que sólo podrá tomar un valor cuando describe a un mismo concepto.

4.3 **Ámbito de especialidad**

El ámbito de especialidad hace referencia al área de especialidad en la que se utilizan un conjunto de lenguajes de especialidad por parte de una determinada comunidad con un

propósito comunicativo determinado. El ámbito de especialidad, además, servirá para conectar el lenguaje con la dimensión comunicativa de la terminología ya que expresa en qué tipo de situación comunicativa podemos utilizar un tipo de lenguaje para hacer referencia a un determinado concepto.

Al indicar el ámbito de especialidad al que hacemos referencia en la descripción de un concepto estamos señalando desde qué punto de vista vamos a describir dicho concepto. Una característica como ‘color’ puede describirse desde numerosos puntos de vista y por eso, indicar formalmente una relación entre nuestra descripción de ‘color’ y su ámbito de especialidad nos permite restringir la visión que estamos tomando para describirla.

Dentro de un ámbito, podemos encontrar conceptos que se catalogan como pertenecientes a distintos lenguajes o sub-lenguajes de especialidad. En el ámbito de la cerámica, encontramos conceptos que provienen del lenguaje de especialidad de la física, la química, la arquitectura, etc. Desde nuestro punto de vista, ocurre que se hace uso de estos lenguajes para dar razón de ser a conceptos del ámbito de la cerámica. Por ejemplo, cuando hablamos de esmalte, en cerámica, no incluimos todos los matices presentes en la descripción de ‘esmalte’ desde el punto de vista meramente químico sino que hacemos uso de ese concepto y de lo que nos interesa de él en nuestro ámbito, su función decorativa, sus propiedades físicas y químicas pero siempre desde el prisma cerámico, es decir, no ahondamos en detalles que puedan interesar a químicos pero no a la industria cerámica.

4.4 **Esencial, no esencial o Distintiva**

Las características mantienen una relación con los conceptos a los que describen que resulta necesario recoger en nuestra descripción de características ya que nos lleva a la tipología que distingue entre características esenciales, características no esenciales y características esenciales distintivas. Esta propiedad señala el grado de relevancia de una característica en la descripción de un concepto en un dominio y da lugar a tres tipos de característica:

- **Característica no esencial.** Cuando una característica no es relevante para la descripción de un concepto porque no añade información que permita identificarlo o distinguirlo de otros conceptos.
- **Característica esencial.** Cuando una característica es relevante para un concepto porque permite identificarlo y describirlo y además, ayuda a distinguirlo de otros conceptos.
- **Característica esencial distintiva.** Cuando una característica es además de relevante la que permite distinguir el matiz que diferencia a un concepto de los demás conceptos hermanos de una clase.

Desde nuestro punto de vista, al indicar si una característica es esencial, no esencial o esencial distintiva estamos describiendo el tipo de relevancia del vínculo que mantiene una característica y el concepto descrito. Cuando se trata de un vínculo que determina la naturaleza del concepto en sí mismo, hablaremos de características esenciales, en cambio, si se trata de un vínculo ocasional que describe el concepto pero no resulta suficiente para explicarlo o no es aplicable a todos los objetos que representan ese concepto por norma general, hablamos de características no esenciales. Finalmente, podemos decir que el vínculo es muy fuerte cuando la característica no sólo servirá para describir el concepto sino además es la que lo diferencia de otros conceptos, la que le hace concepto y no un ejemplar de otro concepto.

Este tipo de vínculo que varía según el concepto al que describe una característica, se puede entender como un tipo de relación que se mantiene entre la característica y un determinado concepto. Entonces hablamos de que existe una relación de tipo esencial, no esencial o distintiva entre una característica y un determinado concepto. Observándolo desde este punto de vista, podemos explicar que **una característica** tenga, **al mismo tiempo**, una relación de tipo esencial con un concepto y **no esencial** con otros, o **distintiva** con un concepto y sólo **esencial** con otros, etc.

Las descripciones que proporcionamos arriba no aclaran cómo juzgar si una característica es relevante o no para un concepto. Por ejemplo, si nos atenemos a aquellas que aparecerían en la definición de baldosa únicamente formarían parte del grupo de las 'características esenciales' cinco o seis características; sin embargo, si nos atenemos a las normas UNE de la Cerámica, el número aumentaría considerablemente. Llegados a este punto conviene establecer qué **criterio** utilizaremos para hablar de características relevantes y no relevantes. Para que una característica se considere esencial deberá cumplir los siguientes requisitos en relación con el concepto al que describe:

- 1- Su **ausencia** podría dar lugar a agrupar el concepto descrito **en clases** a las que no pertenece por naturaleza.
- 2 - Su **presencia** aporta información relevante **para su clasificación** y la clasificación de los conceptos que sean subclase de ese concepto.

De esta forma nos aseguramos una descripción que, desde nuestro punto de vista, parte del concepto de la Terminología y que describirá, mediante la relación de tipo esencial/distintiva que se establece entre una característica y un concepto, un ejemplar ideal de dicho concepto y todas las características necesarias para configurar dicho ejemplar ideal.

A este planteamiento, debemos añadir otro que atañe al hecho de que una característica se constituya de un nombre de característica y un valor. En algunos casos, resulta **esencial que una característica esté presente** en un concepto y su valor varía entre un rango de valores posibles, como vimos a partir de la reflexión del ejemplo de Sager sobre las dimensiones de 'mesa'. En otros casos, resulta **esencial** que la característica esté presente y **posea un valor concreto**, siguiendo con los ejemplos de Sager, la superficie de 'mesa' debe ser 'lisa'. Desde nuestro enfoque, apreciamos que esto sucede también con las características que además de ser esenciales, son distintivas: aunque al reducir la extensión ocurre en menos ocasiones o el rango de valores posibles se reduce. Por ejemplo, una característica distintiva de baldosa respecto a otros tipos de pavimento (parquet, moqueta) es su 'composición' que en este caso es de gres, arcillas, etc. (un rango de valores que engloba minerales y otras materias primas inorgánicas); asimismo, esta propiedad 'composición' también es distintiva de tipos de baldosa y aquí el rango de valores posibles para cada tipo se limita. En resumen:

- a) Podemos observar que las características son esenciales o distintivas porque un concepto y sus hermanos de clase poseen dicha característica (esencial) o porque sus hermanas de clase no poseen dicha característica (distintiva). Por ejemplo, una baldosa tiene una propiedad esencial que es 'anchura' pero el valor concreto de la baldosa no es lo que la caracteriza sino el hecho de que se pueda medir.
- b) Pero también podemos observar que se entiende que una característica es esencial o distintiva porque además de aparecer, su valor es uno en concreto que comparten sus

hermanas de clase (esencial) o que la diferencia de sus hermanas de clase (distintiva). Por ejemplo, una característica esencial de gres rústico es su ‘apariencia’ pero lo que realmente caracteriza a la baldosa es que su ‘apariencia’ sea ‘rústica’.

La representación de toda esta información sobre las características nos ayuda a describir qué relaciones mantiene con los conceptos a los que describe y qué grado de importancia tiene la característica para los conceptos a los que describe. Para ello, necesitamos representar en nuestro trabajo también los conceptos que son descritos por una característica de modo que se pueda establecer la relación de tipo esencial, no esencial o distintivo. Cada uno de los conceptos que es descrito por una característica constituirá el *dominio* de la característica, el “área” conceptual en la que se mueve la característica.

En relación con el dominio, observamos que cuando una característica es esencial para un concepto también lo es para los conceptos a los que agrupa dicho concepto, por ejemplo, si ‘capacidad de absorción de agua’ es esencial para ‘baldosa’ también lo es para ‘gres porcelánico’ (un tipo de baldosa), aunque su rango de valores varíe o sea más limitado. Esto sugiere que en nuestra ontología podremos sacar partido de la representación de estas relaciones para estructurar de forma automática nuevos conceptos que formen parte del dominio de alguna de las características. La estructuración automática se basará en relaciones expresadas anteriormente entre tipos de baldosa y sus características. Si establecemos que un tipo de baldosa B tiene unas características esenciales x, y, z más tarde cuando añadamos un nuevo tipo de baldosa B’ que tiene como características esenciales las mismas x,y,z y además otras, este nuevo tipo de baldosa se considerará un tipo de baldosa de B.

Un ejemplo, si hemos establecido que la baldosa ‘gres porcelánico’ tiene unas características esenciales que son:

- composición: gres y pasta blanca
- capacidad de absorción de agua: $\leq 0.5\%$
- espesor: $> 8\text{mm}$
- abrasión UGL: $110-160\text{m}^3$
- resistencia a la helada: sí
- resistencia química: sí
- carga de rotura: $2200-5200\text{N}$
- medidas: $15 \times 15-60 \times 60\text{cm}$
- acabado: esmaltado

Si nos encontramos con una baldosa de tipo ‘gres porcelánico alto brillo’ que comparte estas características y además tiene otras, como por ejemplo, la característica esencial distintiva ‘acabado: alto brillo’ (entendiendo ‘alto brillo’ un subvalor de ‘esmaltado’), podemos considerar que es un tipo de ‘gres porcelánico’.

En el caso de las distintivas, además de cumplir con lo establecido para ser esenciales tienen un rasgo más. Por eso, cuando una característica es distintiva para un concepto, ésta pasa a ser esencial para los conceptos que pueda agrupar dicho concepto. Por ejemplo, la característica ‘tiene_composición’ con valor ‘gres y pasta blanca’ es distintiva para el concepto ‘gres porcelánico’, y pasa a ser esencial para ‘gres porcelánico alto brillo’ que se podría clasificar como una subclase de la anterior. La diferencia radica que en que las distintivas

además permiten distinguir a un concepto de otros conceptos. Podemos afirmar entonces, que lo que define a una característica distintiva es que cumple dos criterios:

- que la característica es relevante y necesaria para la descripción de un determinado concepto.
- que, en una clasificación determinada, una característica *c* en concreto es la que permite diferenciar al concepto *x* de sus hermanos y su superclase.

Pero, **¿cuántas características distintivas tiene un concepto?** Veamos un ejemplo, si tenemos que describir ‘baldosa_gresificada’ diremos que su característica distintiva respecto a baldosa es ‘tiene_composición’ con valor ‘gres’, si tenemos ‘baldosa rústica’ diremos que su distintiva con respecto a baldosa es ‘tiene_color’ con valor ‘rojo’, estas serán sus características distintivas asumiendo que su superclase es ‘baldosa’. Pero, ¿qué ocurre con ‘baldosa gresificada rústica’? Dependiendo de bajo qué clase o mejor dicho, de entre qué miembros se hace la comparación, la característica distintiva de ‘baldosa gresificada rústica’ es ‘tiene_composición’ con valor ‘gres’ (como subclase de ‘baldosa rústica’), ‘tiene_color’ con valor ‘rojo’ (como subclase de ‘baldosa gresificada’) o incluso la suma de las dos ‘tiene_composición:gres & tiene apariencia:rústica’ (como subclase de ‘baldosa’). Si nuestro enfoque parte de una jerarquización en clases y subclases *a priori* de dominios de las características, podemos organizar la ontología de manera que sólo habrá una característica distintiva en cada uno de los conceptos descritos (véase §3.3.1). Sin embargo, esto supone, primero partir del concepto descrito hacia la característica, aspecto en el que no profundizamos en este trabajo, y segundo, haber razonado, clasificado y pactado la clasificación de los conceptos que vayan a ser descriptos por las características previamente bajo un mismo criterio. Este planteamiento ha sido descartado porque no permite aprovechar las posibilidades de la ontología para representar la **multidimensionalidad** de los conceptos, nos obliga a trabajar en un plano uni-dimensional y además dificulta hacer razonamientos automáticos (inferencia) para ahorrarnos trabajo en el momento de implementar la ontología porque tendremos que haberlo clasificado todo *a priori*. Concluimos entonces que en un plano multidimensional no será una la característica distintiva sino la suma de muchas (todas las posibles en ese contexto) lo que hace distintivo al concepto.

La motivación que nos lleva a concluir esto es que nos interesa describir las características y concretamente en lo que se refiere a si son distintivas o esenciales; ver cómo se jerarquizan de una u otra forma los dominios según la característica que se tenga en cuenta y por eso queremos poder identificar todas las características que se convierten en distintivas o en esenciales para un concepto según los que le rodeen.

4.5 Intrínseca o extrínseca

Las características de los conceptos que son entidades, por ejemplo ‘baldosa’, son intrínsecas cuando se consideran directamente relacionadas con el objeto que representa el concepto y se puede determinar que forman parte de la naturaleza del objeto independientemente de su contexto. Para identificar las características **intrínsecas** de una entidad, podemos **observar el objeto** que lo representa de forma aislada y establecer si podemos detectar esa propiedad **sin** necesidad de otra información ajena al concepto aislado y/o vinculada a nuestra **experiencia** y conocimiento personal para detectarla. En el caso de que para explicar la característica de una entidad necesitemos hacer **uso de información contextual** o basada en la experiencia, estaremos hablando de una característica **extrínseca**.

De acuerdo con esto, consideramos, por ejemplo, que ‘medida’ es una característica intrínseca de ‘baldosa’ y ‘capacidad de absorción de agua’ es una característica extrínseca que viene determinada por nuestro conocimiento sobre el efecto que puede tener el agua en una baldosa.

Una reflexión general sobre los ejemplos y definiciones de característica intrínseca y extrínseca, reseñados en la bibliografía, nos lleva a observar que una misma **característica no puede ser intrínseca para un concepto y extrínseca para otros** y nos sugiere que es algo inherente a la propia característica el ser intrínseca o extrínseca. Desde este enfoque, observamos el carácter intrínseco o extrínseco como una característica de la propia característica, una faceta abstracta que involucra nuestra propia experiencia y conocimiento sobre el mundo.

Dejando a un lado la dificultad que entraña su definición, la representación de esta información sobre las características en la ontología puede ayudar a describir mejor las propias características e incluso aportar información nueva sobre la naturaleza intrínseca o extrínseca de una característica. La información que proporcionará por sí mismo el hecho de describir que una característica es intrínseca o extrínseca es relevante desde la terminología, ya que podemos identificar **qué características**, consideradas de tipo extrínseco, **activan qué relaciones** entre el concepto descrito y otros conceptos y distinguirlas de aquellas que no lo hacen (intrínsecas) pero que se utilizan generalmente como criterio para clasificar conceptos. Además, podemos estudiar qué tipo de característica, intrínseca o extrínseca, **es más relevante** (esencial o distintiva) para la descripción de conceptos, qué características tienen qué tipo de valores como más frecuentes o si es verdad que las características extrínsecas ayudan a explicar que dos conceptos distintos reciban una misma denominación tal y cómo apunta la bibliografía.

4.6 Relación superordinado-subordinado y subordinado-superordinado

Como ocurre con otro tipo de conceptos, las características se pueden estructurar en un sistema conceptual de modo que haya algunas más generales que sirvan para agrupar a otras más específicas. En terminología, esta relación, denominada **superordinado-subordinado** se establece entre un concepto (hiperónimo) que tiene unas características que comparten otros conceptos más específicos (hipónimos), los cuales además tienen otras características.

Esta relación también se establece entre conceptos de tipo característica, como ‘resistencia’, que sirven para agrupar características más específicas, como ‘resistencia a la helada’ o ‘resistencia a la abrasión’. Esta relación superordinado-subordinado que en Ingeniería del conocimiento se suele denominar relación ‘ISA’, considera al concepto subordinado como un tipo de concepto superordinado, es decir, la ‘resistencia a la abrasión’ es un (IS-A) tipo de ‘resistencia’. Asimismo, cuando una característica puede agrupar a otras más específicas, la característica agrupadora se denomina *clase* (de las características agrupadas). En este trabajo, consideramos el uso de esta relación cuando tratamos con nombres de característica que consideramos los representantes lingüísticos de la característica como concepto. En lo que se refiere a los valores, también podríamos identificar este tipo de relación, por ejemplo, se puede hablar de que ‘rojo claro’ es un tipo de ‘rojo’ aunque en este trabajo no vamos a representar la relación superordinado-subordinado entre valores.

Este tipo de relación resulta relevante porque permite estructurar conceptos en forma de jerarquía. En base a esta relación, las características se van organizando en una jerarquía de

forma que podemos entender mejor su naturaleza observando sus hermanos (co-hipónimos) o su 'genealogía' en la jerarquía.

No obstante, en lo que respecta a la jerarquización de conceptos, debemos considerar que el fenómeno de la multidimensionalidad de un concepto puede llevar a distintas clasificaciones para un mismo concepto. Es decir, se puede dar la situación de que una misma característica pueda agruparse bajo dos ramas de la jerarquía diferentes. Este fenómeno puede añadir complejidad a la conceptualización pero al mismo tiempo añade poder descriptivo sobre estos conceptos, ya que nos permite describir un mismo concepto desde más de un punto de vista en un mismo campo de especialidad. Además, remarcar que igual que hay conceptos que se puedan agrupar bajo distintas clases, también es muy común que otros tengan como condición para pertenecer a una clase que no haya otros conceptos en esa misma clase. Esto se puede observar, por ejemplo, entre los conceptos que son características y los que son entidades, si un concepto pertenece a la clase 'característica' no puede pertenecer a la clase 'entidad'.

En una clasificación, el criterio que determina que existe una relación superordinado-subordinado entre uno o más conceptos debe considerar dos aspectos. Por una parte, la clasificación se basará en los objetivos determinados y el tipo de lenguaje que se quiere representar, es decir, si clasificamos conceptos desde el punto de vista de la 'industria cerámica' no debemos basar nuestras relaciones en otro tipo de conocimiento ajeno a este campo de especialidad. Por ejemplo, la característica 'color' es una característica relativa a la 'apariencia' de una baldosa y así se establecerá aunque desde el punto de vista químico puede entenderse como una característica de 'composición' porque constituye el objeto baldosa.

Por otra parte, la clasificación basada en la relación superordinado-subordinado debe ser consistente, para ello podemos aplicar sistemas automáticos, como razonadores, que aseguren la consistencia y nos ayuden a estructurar los conceptos según la información en común que comparten. Por ejemplo, si describimos la característica 'resistencia' e indicamos que sus subordinadas son 'resistencia a la abrasión', 'resistencia a la abrasión superficial' y 'resistencia a la abrasión profunda', cuando añadamos la descripción sobre éstas características, el razonador reordena 'resistencia a la abrasión superficial' y 'resistencia a la abrasión profunda' bajo 'resistencia a la abrasión' que a su vez, seguirá perteneciendo al grupo de 'resistencia'. Sin embargo, además de utilizar este tipo de razonadores, debemos asegurarnos de que no hemos declarado una verdad que no es real, una verdad que la máquina no puede saber que aun siendo coherente es imposible. Por ejemplo, si indicamos que un tipo de 'resistencia' es 'porosidad', 'porosidad' heredará las características de 'resistencia' sea eso verdad o no.

Otro aspecto a considerar es la granularidad de nuestra clasificación, es decir, desde que nivel y hasta que nivel de detalle llegamos en nuestra descripción. Para ello, debemos considerar que en nuestro trabajo, el dominio representa el punto de vista desde el que se representan los conceptos. Por tanto, el nivel superior de la clasificación parte del concepto 'cerámica' y de él nacen los tipos de concepto que vamos a incluir en nuestro trabajo, por ejemplo, las características de la baldosa, objetivo del trabajo, y las baldosas, que se representan para facilitar la descripción de las características, ya que éstas mantienen relaciones importantes con ellas.

Centrándonos en la jerarquización de características de la baldosa, partimos de una primera **clasificación general** que se inspira en las clasificaciones establecidas en la bibliografía sobre terminología y cerámica:

1. **APARIENCIA:** agrupa aquellas características que indican rasgos físicos visibles en la baldosa y que afectan a su estética. Por ejemplo, ‘color’ o ‘textura’.
2. **USO:** representa a la característica que hace referencia a rasgos relacionados con la motivación por la que se hace uso de una baldosa.
3. **COMPOSICIÓN:** representa a la característica que indica rasgos sobre los materiales que conforman una baldosa.
4. **CARACTERÍSTICA FÍSICA:** agrupa aquellas características que indican rasgos internos relativos a la capacidad o rendimiento físico-mecánico de una baldosa. Por ejemplo, ‘módulo de ruptura’ o ‘resistencia a la flexión’.
5. **CARACTERÍSTICA QUÍMICA:** agrupa aquellas características que indican rasgos internos relativos a la capacidad o rendimiento físico-mecánico de una baldosa. Por ejemplo, ‘resistencia a las manchas’ o ‘resistencia a los ácidos y álcalis’.
6. **MEDIDA:** agrupa aquellas características que indican rasgos dimensionales medibles en una característica. Por ejemplo, ‘espesor’ o ‘anchura’.

En el último nivel de descripción, la clasificación se aprovecha de la propia naturaleza de la característica: los valores de una propiedad suponen el último nivel de la clasificación desde el punto de vista teórico, práctico y metodológico.

Estos son los niveles inferior y superior establecidos, a priori, con el fin de permitir que la clasificación tenga un techo desde el que partir y empezar a establecer relaciones superordinado-subordinado con una meta a la que llegar: los valores de la característica.

4.7 Relación de meronimia

Mediante esta denominación describimos la relación que se establece entre una característica en cuyo concepto incluye otra característica sin que ésta última mantenga una relación de hiponimia con la anterior ni tampoco se describa un rasgo. Esta relación la hemos detectado en las características que hacen referencia a rasgos medibles en un concepto que van asociados a unidades de medida. En este tipo de características, observamos cómo puede haber características que miden rasgos de un concepto que son a su vez la suma de otros rasgos identificados mediante otras características. Por ejemplo, el volumen de una baldosa, describe un rasgo que incluye (es resultado de la suma de) otros rasgos identificados mediante las características tamaño y grosor.

Analicemos el caso más representativo que hemos encontrado: la relación que se establece entre la característica ‘planitud de superficie’ y las características ‘alabeo’ y ‘curvatura central’ y ‘curvatura lateral’ a través de algunos contextos.

«Planitud de superficie: Se define en función de medidas obtenidas respecto a tres puntos de la superficie de las baldosas. Las baldosas con relieves en la cara vista que impidan la medición deben ser medidas, cuando sea posible, por la cara posterior.» (Extracto del corpus TXTCERAM, fichero CE023-6e.txt).

De este contexto se desprende que hay tres tipos de medida que determinan el valor de ‘planitud de superficie’.

«Como características dimensionales consideramos la longitud y anchura en piezas rectangulares, el grosor, la rectitud de las aristas o cantos vivos de la pieza, la ortogonalidad del ángulo de los vértices y la planitud de la superficie de la pieza; que incluye la curvatura (central: desviación del centro de la pieza respecto al plano definido por tres de sus vértices; y lateral: desviación del centro de uno de los lados respecto al mismo plano) y el alabeo (desviación del cuarto vértice de una pieza respecto al plano definido por los otros tres).»(Extracto del corpus TXTCERAM, fichero CE023-2e.txt).

Aquí ya podemos ver qué tipos de medida incluye ‘planitud de superficie’, ahora necesitamos observar si estos tipos de medida son tipos de ‘planitud de superficie’ o características que lo describen.

Este extracto de la norma ISO10545-2:1998 (AENOR, 1998a) puede ayudar:

«6 MEDIDA DE LA PLANITUD DE SUPERFICIE (curvatura y alabeo)

6.1 Definiciones

Para los fines de esta parte de la Norma ISO 10545, se aplican las siguientes definiciones:

6.1.1 planitud de superficie: Se define en función de las medidas obtenidas en tres puntos de la superficie de las baldosas.

En el caso de baldosas con relieves en la cara vista que impidan la medición por esta cara, deberán medirse, cuando sea posible, por el reverso.

6.1.2 curvatura central: Es la desviación del centro de una baldosa respecto al plano definido por tres de sus cuatro vértices. (Véase figura 4).

6.1.3 curvatura lateral: Es la desviación del centro de uno de los lados de una baldosa respecto al plano definido por tres de sus cuatro vértices. (Véase figura 5).

6.1.4 alabeo: Es la desviación del cuarto vértice de una baldosa respecto al plano definido por los otros tres. (Véase figura 6).»

Bien, con este pequeño análisis pretendemos resumir que un análisis en profundidad del concepto ‘planitud de superficie’ permite observar que se trata de un concepto que es resultado de la suma de ‘curvatura central’, ‘curvatura lateral’ y ‘alabeo’. Es decir, no podemos afirmar que ‘curvatura central’ es un tipo de ‘planitud de superficie’ sino que es una parte necesaria para que el concepto ‘planitud de superficie’ tenga lugar. Si imaginamos cómo se constituye este concepto mediante conjuntos observamos algo similar a lo siguiente (figura 77):

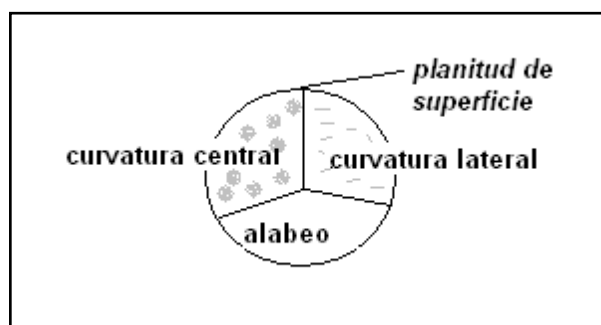


Figura 77. Relación entre planitud de superficie: curvatura central, curvatura lateral y alabeo.

Desde nuestro punto de vista, podemos afirmar que el concepto ‘planitud de superficie’ corresponde a la intersección de los conceptos curvatura y alabeo. Esta relación la hemos detectado al analizar las características de la baldosa y podría entenderse como un tipo de relación parte-todo ya que una característica que constituye un todo incluye otro concepto de tipo característica que resulta necesario para la constitución de la primera aunque desde nuestro

punto de vista resulta arriesgado realizar esta afirmación al carecer de antecedentes bibliográficos que lo corroboren. Sin embargo, sí podemos afirmar que no se trata de un caso aislado, ya que podemos observar y dar importancia a esta misma relación en otros conceptos analizados. Mediante esta relación podemos explicar y describir algunos conceptos de tipo característica que, de otra forma, resultaría más complejo. Por eso, creemos que resulta especialmente interesante formalizar esta relación cuando tratamos con características que aluden a magnitudes físicas complejas o a aspectos mesurables complejos como, por ejemplo, ‘tamaño’, que incluye otros conceptos como ‘altura’ y ‘anchura’ en su intensión.

Respecto a las propiedades matemáticas de la relación, consideramos que puede ser transitiva: un elemento **a** tiene una relación *R* un elemento **b**, un elemento **b** mantiene la misma relación *R* con un elemento **c**, y se puede considerar que el elemento **a** tiene esa misma relación *R* con **c**. Veámoslo con un ejemplo: ‘densidad aparente’ es igual a la ‘masa’ de la baldosa seca partida por su ‘volumen exterior’, podemos decir entonces que ‘densidad aparente’ y ‘volumen exterior’ tienen una relación de tipo característica - característica incluida. A su vez, ‘volumen exterior’ es la suma de la ‘altura’, ‘anchura’ y ‘grosor’ de una baldosa, con lo que podemos decir que ‘volumen exterior’ y ‘altura’ tienen una relación tipo característica – característica exterior. Para que la relación sea transitiva, la relación debe darse también entre ‘densidad aparente’ y ‘altura’, de modo que podamos considerar que ‘altura’ sigue manteniendo el mismo tipo de relación característica – característica incluida, algo que parece darse. Dejando a un lado la necesidad de representar esta relación para según qué conceptos, sí que podemos apreciar que esta relación puede aprovechar la transitividad para mejorar su expresión y que no se incumple esta propiedad. Esta propiedad se cumple también en otras propiedades meronímicas, lo que nos anima a sugerir una vez más, su posible inclusión como relación meronímica.

Por otra parte, la relación de tipo característica – característica incluida no es simétrica, no permite que un elemento **a** tenga una relación *R* con un elemento **b** y a su vez, el mismo elemento **b** tenga la misma relación *R* con **a**. En este aspecto vuelve a resultar similar a otras relaciones meronímicas como componente funcional-objeto, si un concepto **b** está incluido en **a**, no tendría sentido que este se compusiera de **a**. Ocurre lo mismo en lo que refiere a la reflexividad, esta relación no puede ser reflexiva, es decir, el elemento **a** no puede tener la relación *R* con él mismo.

Finalmente, en lo que se refiere a la recursividad, hemos encontrado distintos casos que demuestran que esta relación es recursiva. En los párrafos anteriores hemos presentado ejemplos donde podemos observar que esta propiedad matemática se cumple.

4.8 Relación característica- proceso

En el análisis también hemos encontrado casos donde para describir una **característica** resulta útil explicar su relación con el **método o proceso** utilizado para detectarla o medirla. Esta relación describe el vínculo entre una característica de la baldosa cerámica y el proceso o ensayo realizado para asignar un valor a la característica. En nuestro análisis hemos encontrado numerosos casos donde esta relación resulta relevante para describir a la característica. Por ejemplo, para asignar un valor a la característica ‘densidad relativa aparente’ intervienen dos procesos la ‘impregnación de agua’ (que puede realizarse por ‘método de ebullición’ o ‘método al vacío’) y el ‘pesaje hidrostático’.

Este tipo de relación resulta relevante cuando tratamos con características que expresan rasgos no abstractos de un concepto. Podemos catalogar esta relación como un tipo de relación

argumental similar a *proceso-estado* tal y como se describen la bibliografía (2.3.5.5) donde la actividad determina el estado final del paciente. En este caso, ocurre algo similar a la inversa, se establece una relación entre el estado final del paciente (una característica) y la actividad que determina esa caracterización. Sobre las propiedades matemáticas de esta relación, al igual que la relación proceso-estado, esta relación no es transitiva ni simétrica ni reflexiva y no hemos encontrado casos de recursividad.

4.9 Características de las características

En este apartado presentamos las características que hemos utilizado para describir a las características de la baldosa cerámica.

Las características que hemos identificado en el análisis y que han servido para describir las características de las características de la baldosa cerámica son:

- característica equivalente
- instrumento
- parámetro de ensayo
- unidad de medida

4.9.1 Característica equivalente

Entendemos por características equivalentes aquellas características que aún siendo conceptos distintos describen una misma dimensión de un concepto a través de espacios de valor distintos. Es el caso de ‘capacidad de absorción de agua’ y ‘porosidad aparente’ que hemos explicado anteriormente. Si existe porosidad aparente, significa que la pieza tiene poros que permiten la entrada de aire o líquido, la capacidad de absorción de agua mide la capacidad de una pieza de permitir la entrada de agua, a través de los poros. Siendo conceptos distintos, hacen referencia a la misma dimensión en la descripción de un concepto.

En este aspecto, la relación causa-efecto, tal y como se explica en la bibliografía (véase II.2.3.5.6) y aplicada a las características, no difiere mucho del fenómeno que considera a dos características equivalentes (véase §II.2.2.6). La relación **causa-efecto** tiene lugar cuando, por ejemplo, una característica implica otra esté presente en una misma entidad. Nuevamente, depende de la perspectiva que tomemos la consideración de ‘porosidad aparente’ y ‘capacidad de absorción de agua’ como características equivalentes o como características que mantienen una relación causa-efecto. No obstante, esta información resulta relevante, independientemente de cómo se la trate, y por eso en este trabajo la hemos considerado limitando que la equivalencia existe cuando las características describen a un concepto y no como conceptos en sí mismas, donde mantienen una extensión e intensión distintas.

4.9.2 Instrumento de ensayo

En la descripción de las características de la baldosa cerámica encontramos numerosas referencias al instrumento utilizado para medir o ensayar la característica en la baldosa. Esta característica servirá para describir los rasgos que distinguen o comparte una característica con sus hermanas de clase y para aportar información sobre la intensión y la extensión de la característica. Por ejemplo, saber que ‘resistencia al rayado’ “tiene_instrumento: mineral” nos ayuda a discernir que la resistencia al rayado se refiere a algo físico, probablemente relativo a la

capacidad de la superficie de no dañarse al arrastrar un material (mineral) sobre ella y descarta la posibilidad de que haga referencia a si le pueden dibujar rayas.

4.9.3 **Parámetro de ensayo**

El parámetro de ensayo es otra característica que nos permite conocer más acerca de la intensidad y la extensión de una característica. Para determinar los valores de una característica, existen casos en los que durante el proceso de medida de la característica se utilizan unos determinados parámetros establecidos que bien pueden ser fórmulas acordadas para realizar los cálculos, características a tener en cuenta en la apreciación del valor u otro tipo de dato. Por ejemplo, para determinar si existen o no existen pequeñas diferencias de color en una baldosa, entran en juego el *factor comercial*, el valor de tolerancia acordado en la industria de baldosas cerámicas para determinar la aceptabilidad de la diferencia de color a través de fórmulas o acuerdos entre la comunidad cerámica; o los valores del espacio de color CieLab, valores calculados a partir de las curvas de reflexión espectral catalogados por la *Commission internationale de l'éclairage*. Esta información resulta relevante, ya que desde el punto de vista del terminólogo interesa saber qué criterio se utiliza para determinar los valores de una característica, por ejemplo, los valores de 'resistencia al deslizamiento' se determinan en base a los resultados del 'coeficiente de fricción'. Otro ejemplo de esta característica para 'pequeñas diferencias de color' es "tiene_parámetro: factor comercial, espacio de color CieLab, cromaticidad, luminosidad".

4.9.4 **Unidad de medida**

Otra característica que vamos a representar en nuestra ontología es la que indica la unidad de medida utilizada para expresar los valores de una característica. Esta característica proporciona información sobre la intensidad de la característica al expresar en sí misma que será una característica cuantificable según un estándar y de extensión al restringir los valores (vistos como ejemplares de la característica). Hemos encontrado numerosos ejemplos en el análisis del ámbito, por ejemplo: 'anchura' "tiene unidad de medida: centímetro". Esta información resulta relevante, ya que desde el punto de vista del terminólogo a menudo interesa más saber qué se mide y cuál es la unidad de medida que se utiliza y no tanto cuál es el valor numérico en concreto.

4.10 **Sinonimia y equivalencia en otra lengua**

La representación de la relación de **sinonimia** nos permite aumentar la expresividad de la ontología y permite que el usuario pueda acceder a la información desde distintas denominaciones de un mismo concepto. En nuestro trabajo, esta relación se puede observar entre dos o más denominaciones de una misma característica o bien, entre dos o más denominaciones de valores de una característica en una misma lengua o entre lenguas. Por ejemplo, 'planitud' y 'planitud de superficie' en el contexto de la cerámica son ejemplos de denominaciones sinónimas que aluden a una misma característica conceptual.

Esta relación, tal y como se ha explicado anteriormente, puede tener como propiedades matemáticas la recursividad, la simetría y la transitividad (§II.2.3.6.4) y también puede observarse en otro tipo de conceptos de la ontología como entidades o actividades.

Es necesario remarcar que el objetivo de este trabajo no pasa por elaborar una ontología bilingüe a todos los efectos sino probar las posibilidades de formalización de la ontología de manera que sea independiente del lenguaje, de ahí que nos centremos en la recogida de algunos

equivalentes en inglés para las características de la baldosa cerámica pero no profundizamos en la extracción de equivalentes de manera exhaustiva.

4.11 Otra información en lenguaje natural

Nuestro propósito principal es formalizar el máximo de información posible sobre las características de la baldosa de manera que se pueda procesar y reutilizar mediante un sistema informático. No obstante, consideramos de utilidad incluir otro tipo de información contextual o explicativa, en lenguaje natural, que ayude al usuario humano a entender la ontología y los elementos que la componen.

En el caso de las características, hemos recogido contextos donde aparece la característica y la fuente de los contextos. Entre la información que añadiremos sobre las características, se incluyen también referencias a la/s fuente/s utilizada/s para extraer información sobre la característica, explicitando la norma o documento donde se puede consultar más información sobre la característica.

Otro tipo de explicaciones recogidas en la ontología, que se pueden observar tanto en características como en otros conceptos, se refieren a información sobre aspectos que el usuario de la ontología debe considerar a la hora de utilizar una determinada característica o describir un determinado concepto en la ontología. Por ejemplo, durante la fase de extracción de información detectamos que los valores de 'acabado superficial' podían expresarse en porcentaje de baldosas libres de efectos o en valores booleanos: 'cumple la norma' (si al menos el 95% de las baldosas están libres de efectos) o 'no cumple la norma'. Dado que en la formalización de la característica se va a priorizar la expresión de valores en una de las dos formas, conviene recoger esta información para que el usuario sea conocedor de que existe otra alternativa.

4.12 Propuesta de modelo de descripción de una característica

A partir del análisis de la bibliografía y contrastando esta información con el corpus, hemos estructurado e identificado los elementos que pueden ayudar a describir una característica. En este análisis hemos identificado los rasgos que permiten ordenar y organizar esta información sobre las características, algunos de ellos discutibles por su novedad y falta de referencias pero que se justifican en la necesidad de expresar y describir el concepto de característica.

A continuación, hemos elaborado una representación del modelo de descripción de la característica. Este modelo (figura 78) incluye todos los conceptos que participan en la descripción de una característica y aparecen vinculados entre ellos según deben considerarse para la formalización en la ontología. Como podrá observar el lector, la información relativa al rol de la característica en la descripción de un concepto se ha representado a parte, en un recuadro inferior. Las razones que nos han llevado a tomar esta decisión es que en un plano unidimensional resultaba complicado espacialmente representarlo todo sin crear confusión en la figura. La finalidad de esta representación es ayudar al diseño de la ontología, observando qué tipo de conceptos deben vincularse y cómo organizamos esta información sobre la característica.

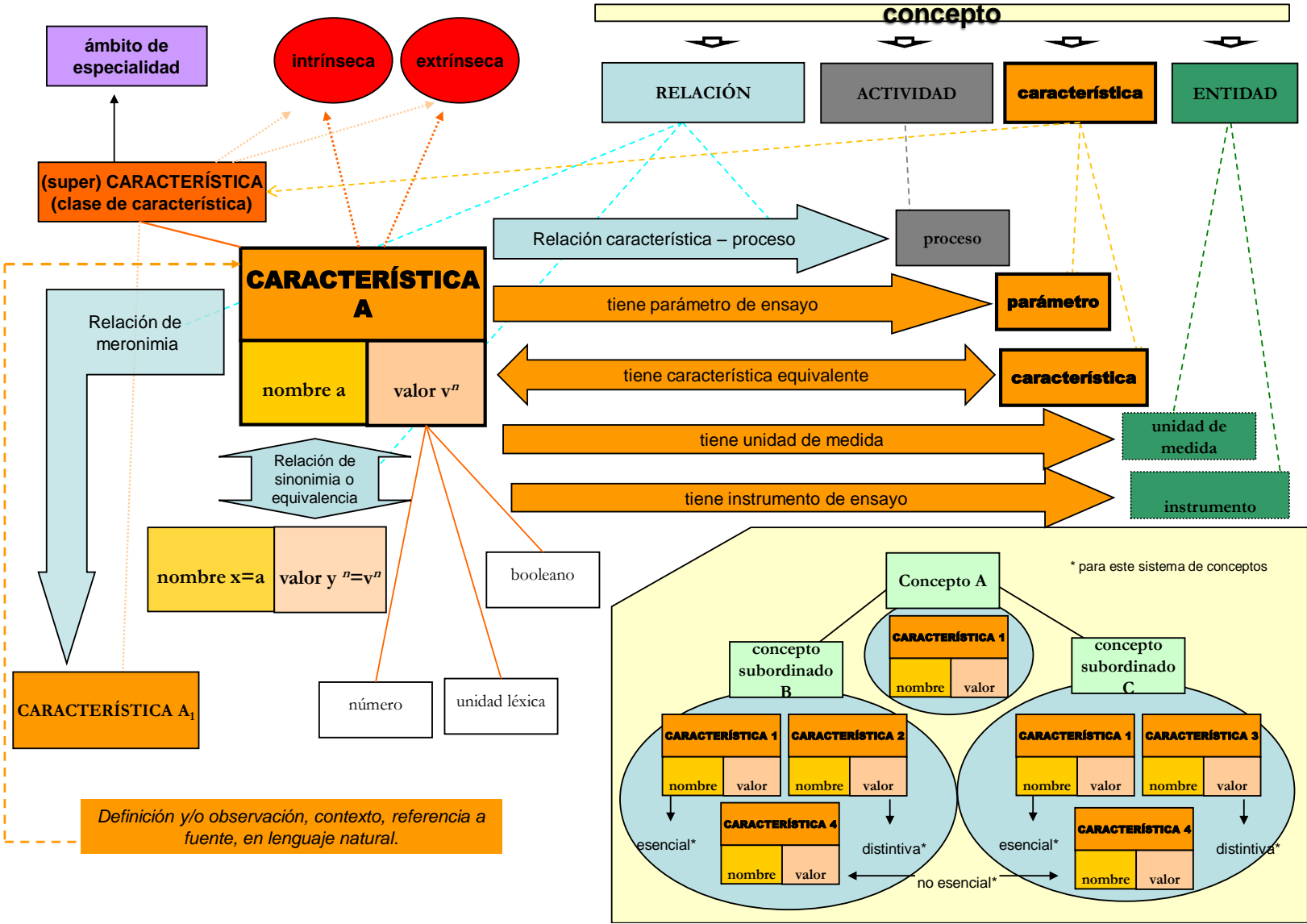


Figura 78. Modelo estático de descripción de una característica y su rol en la descripción de conceptos.

En el centro de la figura 78, observamos el concepto ‘característica’ que se considera el núcleo del modelo alrededor del cual se van estableciendo vínculos con otros elementos que ayudan a describirla. La mayoría de elementos se organizan según el tipo de concepto que constituyen, salvo el elemento ‘explicación en lenguaje natural’ de la característica y el ‘ámbito’ que se han representado de forma independiente: el primero, porque no es relevante considerar su tipología y el segundo, porque su tipología es variable y pretendemos que el modelo resulte generalizable.

En la representación gráfica, hemos utilizado los colores cálidos (rojo, naranja, calabaza, arena) para representar elementos relativos exclusivamente al concepto ‘característica’. El color azul claro se ha utilizado para representar los elementos que son conceptos de tipo ‘relación’ y hacen referencia a las posibles relaciones que mantiene una característica con otras o con otros conceptos. Los conceptos de tipo ‘entidad’ aparecen representados en color verde y el color gris oscuro hace referencia a conceptos de tipo ‘actividad’. Hemos representado la jerarquía que se establece entre estos elementos, utilizando también esos colores, por ejemplo, el naranja si el vínculo es entre elementos que forman parte del concepto característica (o son tipos de característica) o el verde para indicar que un elemento concreto es un tipo de entidad etc.

En la parte superior central de la figura se ha representado la característica compuesta del elemento nombre y elemento valor. El elemento valor agrupa a los tres tipos de valores que vamos a considerar: valores de tipo booleano, valores de tipo numérico y valores que son unidades léxicas. El concepto característica, además, aparece supeditado por su clase, que es a su vez otro concepto de tipo característica. Tanto la clase como la característica se considerarán de tipo extrínseco o intrínseco, por eso, hemos representado con una flecha discontinua la posibilidad de pertenecer a uno de estos dos grupos. Finalmente, otro concepto supeditado a la característica es su ámbito que hemos representado indicando que la clase de la característica pertenece a éste a través de una flecha.

Las relaciones conceptuales se han representado mediante flechas con texto que indican la dirección de la relación y que se sitúan entre los conceptos que mantienen la relación. Si las recorremos en la figura siguiendo el orden contrario a las agujas del reloj, encontramos:

- Relación de meronimia, entre dos características
- Relación de sinonimia, entre dos o más denominaciones
- Relación característica – proceso de ensayo, entre una característica y uno o más procesos relacionados con su ensayo

Las características de las características aparecen en color anaranjado con forma de flecha. Se han expresado inspirándose en la representación de características en forma de atributo-valor. Están situadas en el centro de la figura y por orden de aparición en vertical son:

- Tiene parámetro – parámetro, describe el parámetro de ensayo practicado para determinar los valores de una característica
- Tiene característica equivalente – característica, describe la equivalencia de la dimensión de un concepto que describen dos o más características
- Tiene unidad de medida – unidad de medida, describe la unidad de medida utilizada para expresar los valores de una característica
- Tiene instrumento de ensayo – instrumento, describe el instrumento o instrumentos utilizados para ensayar los valores de una característica

En la parte inferior derecha de la figura, hemos querido reflejar el rol de la característica que describe a un concepto y su relevancia para éste según el tipo de información que aporta. No es una representación exhaustiva, pero sí puede ayudar a comprender qué distribución tomará una característica en un conjunto de conceptos para que sea considerada esencial, distintiva o no esencial. A través del ejemplo, observamos que la característica 1 describe al concepto A que subordina al concepto subordinado B y el concepto subordinado C, éstos conceptos subordinados también tienen esa característica 1 y el hecho de que la compartan con su clase y entre ellos, viene a reflejar que se trata de una característica esencial para pertenecer a ese grupo. Al mismo tiempo el concepto subordinado B se distingue del C porque tiene la característica 2, con lo que podemos decir que la característica 2 es distintiva. Igual que ocurre con la característica 3 para el concepto C. Finalmente, la no relevancia la hemos reflejado mediante la característica 4 que tanto B como C tienen pero no resulta relevante para pertenecer a A ni para distinguirse entre ellos.

Resulta evidente, sin embargo, que la representación gráfica del modelo de descripción de una característica en un plano unidimensional y de forma estática no nos ha bastado para aglutinar todos los rasgos que queremos representar en la ontología. Por ejemplo, no hemos representado que una característica no puede ser de tipo intrínseco y extrínseco al mismo tiempo porque no hemos encontrado una forma que resulte clara, expresiva y que ayude a realizar su posterior formalización en la ontología sin acabar creando una figura compleja y difícil de leer. Por eso, además de la información que representa la figura, hemos elaborado un listado de requerimientos y premisas necesarios para representar formalmente el concepto de característica en nuestro trabajo.

Este paso es muy relevante porque determina y delimita el modelo formal y las reglas y datos que debe contener la ontología. Entre los requerimientos, distinguimos entre especificaciones generales y especificaciones de concepto. Las especificaciones generales hacen referencia a aspectos teóricos referentes a enfoques de formalización y representación de conceptos y metodologías que vamos a utilizar para formalizar las características en la ontología. Las especificaciones de concepto también están vinculadas a aspectos teóricos pero se centran en aspectos puntuales relativos a un tipo de concepto o criterio a considerar en la descripción de un determinado rasgo en la ontología.

Las especificaciones generales son:

- La formalización de los conceptos debe ser independiente del lenguaje de manera que pueda agrupar otras denominaciones o equivalentes en otras lenguas sin modificar la estructura.
- La delimitación de características distintivas debe basarse en la diferencia única, es decir, los miembros de una subclase deben distinguirse entre ellos a partir de una misma característica o unas mismas características.
- La formalización debe ser lo suficientemente flexible como para permitir la herencia múltiple de clases. Es decir, una clase debe formalizarse de manera que pueda pertenecer a tantas clases como sea posible de acuerdo con su concepto.
- La formalización debe ser consistente y tener el detalle suficiente y necesario como para que una clase se agrupe a una superclase basándose en criterios de comparación de las características y las relaciones de las clases y no sólo por asignación directa a la clase.

Las especificaciones de concepto son:

- Una **característica** es un tipo de concepto distinto a una entidad, una actividad y una relación que puede representarse como concepto y/o como elemento descriptor de conceptos en forma de dominio-nombre de característica-valor.
 - A efectos de computación, la representación de una característica en forma de dominio-atributo-valor, no resulta diferente de la representación de una relación en forma de dominio-nombre de relación-valor.
- Las características tendrán características que las describan.
- Una característica puede tener una relación de **subordinación** con otra característica más general o de **superordenación** con otra característica más específica. La característica más general que puede englobar otras más específicas se considera la **clase** de las características más específicas. Una característica puede pertenecer a más de una **clase**.
- La relevancia de una característica para la descripción de un concepto determina que la característica se considere **esencial, distintiva o no esencial**.
 - Una característica se considera esencial cuando resulta relevante para la descripción de un concepto.
 - Si además de relevante sirve para diferenciar a un concepto de otros de su misma clase, la característica se considera distintiva.
 - Una característica se considera no esencial cuando su papel en la descripción de un concepto no es decisivo para explicar el concepto. Las características no esenciales no se recogen en la ontología.
- El **nombre de la característica** sirve para **agrupar** los **valores** de la característica. Un **valor** indica un **rasgo concreto** de un concepto.
- Se distingue entre tres tipos de valor: una **unidad léxica**, un número **porcentual** o un **número acompañado** de una **unidad de medida** y un valor **booleano**.
 - El rango de **valores** de tipo **unidad léxica y de tipo numérico** de una característica puede ser **limitado** o **ilimitado** dependiendo de la característica y/o del rango aceptado que pueda tomar cuando describa a un determinado concepto.
- Una característica puede ser **intrínseca** o **extrínseca**. Las características **intrínsecas** describen al concepto desde su propia naturaleza. Las características **extrínsecas** describen al concepto a partir de la relación de éste con otros conceptos y requieren un conocimiento añadido basado en la experiencia.
 - Una característica no puede ser **intrínseca** y **extrínseca** al mismo tiempo.
- Una característica pertenece siempre a un **ámbito de especialidad** y éste debe coincidir con el ámbito al que pertenece su **clase**.

5 Extracción de las características de la baldosa cerámica y su información conceptual

Para llevar a cabo la extracción de características de la baldosa hemos utilizado procedimientos y recursos distintos (ver descripción detallada de fuentes en “Recursos y herramientas”). Esta etapa se ha realizado en tres fases:

- Extracción de una lista inicial de características de la baldosa cerámica.
- Extracción de información conceptual sobre las características y verificación de la lista de características.
- Extracción de ejemplos de tipos de baldosa y descripciones de baldosas
- Validación de la información con un experto

5.1 Extracción de una lista inicial de candidatos a característica del producto acabado.

La lista inicial de características del producto acabado cerámica se ha extraído a partir del análisis de las normas UNE-EN 14411 y 10545 (§1.1). La utilización de estas normas de estandarización nos ha permitido identificar de forma precisa y concisa las características que se utilizan en el ámbito cerámico para clasificar y describir las baldosas. La estructura de las normas y su contenido nos ha permitido extraer, además de un listado de características, otra información relevante como relaciones entre características y/u otros conceptos, valores de características, métodos para obtener o medir estos valores o los tipos de baldosa que describen y los criterios utilizados, información sobre la relevancia de una características sobre otros para según qué baldosas, etc.

La lectura y análisis de estas normas ha dado como resultado la extracción manual de 38 características de la baldosa cerámica:

- 1) alabeo
- 2) anchura
- 3) aspecto superficial
- 4) capacidad de absorción de agua
- 5) carga de rotura
- 6) cromaticidad
- 7) curvatura central
- 8) curvatura lateral
- 9) densidad aparente
- 10) densidad relativa aparente
- 11) dilatación por humedad
- 12) dilatación térmica lineal
- 13) dimensiones
- 14) emisión de cadmio y plomo
- 15) espesor
- 16) fuerza de rotura
- 17) longitud
- 18) luminosidad
- 19) ortogonalidad

- 20) pequeñas diferencias de color
- 21) piezas especiales
- 22) planitud de superficie
- 23) porosidad aparente
- 24) rectitud de lados
- 25) resistencia a ácidos y álcalis
- 26) resistencia a ácidos y álcalis en concentraciones débiles
- 27) resistencia a ácidos y álcalis en concentraciones fuertes
- 28) resistencia a la abrasión profunda
- 29) resistencia a la abrasión superficial
- 30) resistencia a la flexión
- 31) resistencia a la helada
- 32) resistencia a las manchas
- 33) resistencia a los productos domésticos de limpieza y aditivos para agua de piscinas
- 34) resistencia al choque térmico
- 35) resistencia al cuarteo
- 36) resistencia al deslizamiento
- 37) resistencia al impacto
- 38) resistencia química

Asimismo, esta lista se ha completado con 4 características encontradas en los catálogos de cerámica analizados: acabado, color, textura y uso. La identificación de estas características también se ha llevado a cabo de forma manual mediante la lectura y análisis de los catálogos disponibles. El listado de características para la extracción de información se basa en estas 42 características.

5.2 Extracción de información conceptual sobre las características.

En esta fase de extracción de información sobre las características, el objetivo era recoger toda la información necesaria para describir una característica basándonos en los criterios que habíamos establecido en la fase de análisis de la bibliografía.

Para llevar a cabo la recogida de información hemos elaborado una serie de documentos para la recogida de datos:

- Tabla de descripción general de características (anexo 1)
- Listado de instrumentos (anexo 3)
- Listado de parámetros (anexo 4)
- Listado de procesos (anexo 5)
- Listado de unidades de medida (anexo 6)
- Glosario de equivalencias en español e inglés de las características y baldosas (anexo 7)

En la tabla de descripción general de las características, hemos recogido información sobre las características. La tabla que presentamos en el anexo 1 incluye el listado de características con las que hemos trabajado y contiene los siguientes campos de información acerca de éstas:

- | | |
|----------------------------|-------------------------------|
| - Número de identificación | - Otras denominaciones |
| - Fuente | - Valores o ejemplos de valor |

- | | |
|--------------------------------------------------------------------------------------------|--------------------------------|
| - Tipo de valor | - Características equivalentes |
| - Especificación del tipo de rango (limitado o ilimitado) y si procede, valores del rango. | - Unidad de medida |
| - Categoría general | - Parámetro de ensayo |
| - Hiperónimos | - Relación con proceso |
| - Dominio de la característica | - Meronimia |
| - Característica intrínseca o extrínseca | - Contexto 1 |
| - Instrumentos | - Contexto 2 |
| | - Contexto 3 |
| | - Ámbito |

El contenido de cada campo se describe a continuación:

- Número de identificación – contiene un número que se ha utilizado para facilitar tareas de gestión interna de datos.
- Denominación de la característica – este campo contiene la denominación principal de la característica, el resto de denominaciones aparecerán en un campo a parte.
- Fuente de la característica/ Fuente complementaria – en este campo indicamos si la característica la referencia a la fuente de la característica, indicando la norma en concreto o el catálogo comercial donde aparece. Igualmente, puede contener una referencia a una fuente complementaria que hayamos detectado durante el proceso de análisis.
- Otras denominaciones – este campo incluirá el resto de denominaciones de la característica encontradas en los recursos textuales.
- Valores y/o ejemplos de valores – este campo contiene los valores posibles de la característica o en el caso de que estos valores sean muy numerosos, ejemplos de valores para poder identificar su estructura. Por ejemplo, en el caso de valores numéricos seguidos de unidades de medida, se recogen algunas muestras encontradas en los recursos textuales utilizados.
- Tipo de valor – este campo describe el tipo de valor o tipos de valor que corresponden a la característica. La información de este campo se ha homogeneizado y se limita a los siguientes tipos: conceptos presentes en la ontología, booleano, *double*, *integer*.
- Hiperónimos – este campo incluye otras clases bajo las cuáles podemos agrupar a la característica y que hemos obtenido del análisis de los recursos textuales
- Dominio – este campo contiene el tipo de baldosa que describe la característica. Los valores posibles son baldosa, baldosa esmaltada y baldosa no esmaltada.
- Intrínseca/extrínseca – en este campo indicamos si la característica es intrínseca o extrínseca
- Instrumentos – este campo incluye el instrumento o instrumentos que se utilizan para ensayar los valores de la característica
- Característica equivalente – este campo incluye la característica equivalente a la característica que se describe
- Unidad de medida – este campo incluye la unidad de medida que mantiene la relación con la característica
- Parámetro – este campo incluye el parámetro o parámetros utilizados para ensayar los valores de la característica

- Relación con proceso de ensayo – este campo incluye el proceso o procesos utilizados para determinar los valores de la característica
- Relación de meronimia – este campo incluye la denominación de la característica que actúa como parte con respecto a la característica analizada
 1. Contexto 1– este campo contiene contextos donde aparece la característica extraídos del corpus TxtCeram y la referencia al documento del corpus.
 2. Contexto 2 – este campo contiene contextos donde aparece la característica extraídos de las normas UNE de cerámica (UNE-EN ISO 10545 o UNE-EN 14411) y la referencia a la norma o parte de la norma en concreto.
 3. Contexto 3 – este campo contiene contextos donde aparece la característica extraídos de los catálogos comerciales de cerámica y la referencia al en concreto.
 4. Área temática – este campo hace referencia a la cerámica en todos los casos.

De forma paralela, hemos recogido en sendos documentos: una lista de instrumentos de ensayo (anexo 3), una lista estructurada de parámetros de ensayo (anexo 4), una lista de procesos de ensayo (anexo 5) y una lista de unidades de medida de las características (anexo 6). A medida que se iban identificando los instrumentos, parámetros, procesos o unidades de medida en el estudio de una característica, cada uno de estos conceptos se incluía en un documento específico para cada tipo de dato y se complementaba con observaciones o comentarios acerca de su relación con las características o una explicación del concepto. Estas listas de conceptos por tipos se utilizarán más tarde para crear las clases en la ontología. Los listados finales se han elaborado en formato de texto, utilizando la tabulación para estructurar las clases y subclases de instrumento, proceso o parámetro identificadas en los textos. La estructura de cada listado se basa únicamente en las clasificaciones identificadas en los documentos, es decir, no ha habido un trabajo de reagrupación o clasificación por nuestra parte. Este tipo de conceptos nos interesan con fines de descripción de la característica y por tanto no se ha realizado un análisis suficientemente exhaustivo como para proponer una jerarquía de clases de instrumentos, parámetros o procesos más allá de la que hemos observado a partir del análisis de los documentos.

La tabla y los listados son el resultado final de un proceso de reelaboración continuado. Respecto a la tabla, a lo largo del proceso de extracción de datos y comparación con el análisis de la bibliografía hemos ido añadiendo nuevos campos a medida que íbamos detectando nuevos tipos de información que podría resultar relevante y hemos revisado cada entrada para ajustarla a las nuevas decisiones. Por ejemplo, notamos que resultaría interesante afinar el detalle del campo que recogía el tipo de valor de una característica, en principio distinguíamos entre léxico, valor booleano y número, pero más adelante detectamos que había características cuyos valores numéricos incluían números decimales y otras que únicamente incluían enteros. Esto nos llevó a redefinir los posibles valores del campo en la tabla.

A través del análisis de los recursos textuales, hemos completado los campos de la tabla intentando que no quede ninguno vacío. Sin embargo, dependiendo de la naturaleza de la característica algunos campos no contienen información porque, por ejemplo, la característica no mantiene un determinado tipo de relación o no tiene una unidad de medida.

Al mismo tiempo que se recogía información para la tabla de datos de las normas UNE de cerámica, hemos cruzado el listado con el corpus TxtCeram y con los catálogos comerciales de cerámica. En el caso de la búsqueda en el corpus, el proceso ha consistido en buscar una por una

cada característica del listado y utilizar la información proporcionada en cada búsqueda para completar los apartados de la tabla de datos: contextos (si aportaban información nueva), hipónimos/hiperónimos, información relevante, valores, etc.

La herramienta que hemos utilizado durante este proceso es Concord de WordSmith. Esta herramienta nos ha proporcionado listados de ocurrencias en contextos de 7 palabras por delante y por detrás. En el caso de los catálogos y las normas que estaban en formato pdf, los contextos se han extraído manualmente, utilizando las opciones de búsqueda avanzada de Acrobat Reader. Los catálogos han resultado de gran utilidad para recoger los valores de las características y agregar información relativa a su dominio (tipos de baldosa que suelen describir).

Asimismo, hemos consultado las entradas de la lista en el Diccionario Cerámico Científico-Práctico (Guillem Monzonís y Guillem Villar, 1987). Este proceso ha sido casi completamente manual y se ha realizado a partir de un documento electrónico que contiene todas las entradas del diccionario digitalizado para así permitir el uso de buscadores sencillos de palabras.

Finalmente, otro documento de trabajo elaborado es un pequeño glosario español-inglés de las características de la baldosa cerámica que hemos trabajado. La extracción de equivalentes al inglés de las características se ha realizado comparando las norma UNE-EN 14411 con la norma ISO equivalente en inglés (ISO, 1998) y las partes de la norma UNE-EN ISO 10545 con las ISO homólogas en inglés (ISO, 1995-1999) o consultado los catálogos comerciales que contenían equivalentes a otras lenguas. El resultado de esta extracción no persigue la elaboración un glosario exhaustivo de equivalentes en inglés de cada concepto de la ontología sino obtener una muestra de equivalentes para probar las posibilidades de representación de conceptos de manera independiente del lenguaje.

5.3 Recopilación de descripciones sobre tipos de baldosas

En una tercera fase de trabajo hemos recogido las descripciones de tipos de baldosa que figuran en la norma UNE-EN 14411 de 2007 (AENOR, 2007) y las descripciones de tipos de baldosa presentes en el capítulo 1 del manual de PROALSO (PROALSO, 2011).

La finalidad de esta tarea no es tanto incluir una lista exhaustiva de todos los tipos de baldosa sino una muestra relevante que nos sirva para probar la adecuación de las características elegidas y del planteamiento utilizado para representarlas en la ontología, especialmente en relación con su papel como descriptoras y diferenciadoras de tipos de baldosa.

Para recoger la información sobre las baldosas, hemos creado dos fichas, una para los tipos de baldosa según la norma UNE 14411 y otra para los tipos de baldosa identificados en el manual de PROALSO. Cada ficha de descripción de tipos de baldosa incluye la siguiente información:

- 1) Número de identificación – hemos numerado los tipos de baldosa para facilitar tareas de gestión interna de datos.
- 2) Denominación de la baldosa – este campo contiene la denominación principal de la baldosa, el resto de denominaciones se recogen en el anexo 7.
- 3) Fuente de la baldosa – en este campo recogemos la fuente dónde hemos localizado la denominación de la baldosa.

- 4) Características y valores – cada una de las características se presenta en un campo y su valor a aparece en el campo contiguo.

Las fichas se han elaborado utilizando Ms Excel. En cada ficha, cada una de las características identificadas en alguna de las baldosas se ha introducido como encabezado de una columna. El valor de cada característica para cada baldosa se ha introducido en la celda correspondiente a su columna y fila. Si un tipo de baldosa no tiene una determinada característica la celda correspondiente a su fila se deja vacía. De esta forma, podemos utilizar filtros que nos permitirán identificar rápidamente las características distintivas o esenciales según grupos de baldosas y tener en cuenta esta información para formalizarla en la ontología.

5.3.1 Clasificación de la norma UNE-EN 14411

En la norma UNE-EN 14411 (AENOR, 2007), encontramos una clasificación de baldosas según su absorción de agua y método de moldeo de la que hemos extraído tipos de baldosa e información sobre dos de sus características. En la tabla 22, se muestra la tabla de clasificación que proporciona la norma UNE-EN 14411 de 2007:

Clasificación de las baldosas cerámicas según su absorción de agua y método de moldeo

Moldeo	Grupo I $E \leq 3\%$	Grupo II _a $3\% < E \leq 6\%$	Grupo II _b $6\% < E \leq 10\%$	Grupo III $E > 10\%$
A Extruidas	Grupo AI (véase anexo A)	Grupo AII _{a-1} ^a (véase anexo B)	Grupo AII _{b-1} ^a (véase anexo D)	Grupo AIII (véase anexo F)
		Grupo AII _{a-2} ^a (véase anexo C)	Grupo AII _{b-2} ^a (véase anexo E)	
B Prensadas en seco	Grupo BI _a $E \leq 0,5\%$ (véase anexo G)	Grupo BII _a (véase anexo J)	Grupo BII _b (véase anexo K)	Grupo BIII ^b (véase anexo L)
	Grupo BI _b $0,5\% < E \leq 3\%$ (véase anexo H)			
C Fabricadas por otros métodos	Grupo CI ^c	Grupo CII _a ^c	Grupo CII _b ^c	Grupo CIII ^c

^a Los grupos AII_a y AII_b se subdividen en dos partes (1 y 2) con diferentes especificaciones de producto.
^b El grupo BIII incluye solo a las baldosas esmaltadas. Hay una pequeña cantidad de baldosas prensadas en seco, no esmaltadas con absorción de agua mayor de 10% que no se incluyen en este grupo de producto.
^c Estas baldosas no se incluyen en esta norma europea.

Tabla 22. Tabla de clasificación de baldosas cerámicas según la norma UNE-EN 14411 de 2007 (AENOR, 2007)

Las clases de baldosa representadas en esta tabla son: por una parte, baldosa extruida, baldosa prensada en seco y por otro, baldosas del grupo AI, grupo BI_a, grupo BI_b, grupo BI_a, grupo CI, grupo AII_{a-1}, grupo AII_{a-2}, grupo AII_{b-1}, grupo AII_{b-2}, grupo BII_b, grupo CII_b, grupo AIII, grupo BIII y grupo CIII.

Además, a partir de los criterios de la primera tabla de clasificación de baldosas, la norma UNE-EN 14411 de 2007 incluye una descripción más detallada sobre cada grupo de baldosas incluyendo otras características que resultan necesarias para las baldosas de cada uno de los grupos. En la tabla 23, podemos observar parte de la tabla de descripción de baldosas del grupo AI.

Este tipo de tablas se han utilizado para completar las fichas de descripción de las baldosas de modo que se ha ido agregando información sobre las distintas clases. Al mismo tiempo, este tipo de tablas junto con las que presentamos a continuación, nos han servido para

identificar las características esenciales y distintivas para las distintas clases de baldosa, mediante el análisis y la comparación de características y valores.

Requisitos para baldosas extruidas, Grupo AI, $E \leq 3\%$

Dimensiones y aspecto superficial	Precisión	Natural	Ensayo
Longitud y anchura			
El fabricante debe elegir las dimensiones de fabricación de forma que:			
a) permitan una junta de anchura nominal entre 3 y 11 mm ^a para baldosas modulares;			
b) la diferencia entre la medida de fabricación y la medida nominal no sea mayor de ± 3 mm para baldosas no modulares.			
Desviación admisible en % de la medida media de cada baldosa (2 ó 4 lados) respecto de la dimensión de fabricación (<i>W</i>)	$\pm 1,0\%$ hasta un máximo de ± 2 mm	$\pm 2,0\%$ hasta un máximo de ± 4 mm	ISO 10545-2
Desviación admisible en % de la medida media de cada baldosa (2 ó 4 lados) respecto de la medida media de las 10 probetas (20 ó 40 lados)	$\pm 1,0\%$	$\pm 1,5\%$	ISO 10545-2
Espesor			
a) el fabricante debe especificar el espesor			
b) desviación admisible en % del espesor medio de cada baldosa respecto a la dimensión de fabricación correspondiente.	$\pm 10\%$	$\pm 10\%$	ISO 10545-2
Rectitud de lados^b (cara vista)			
Máxima desviación de rectitud en % con relación a la dimensión de fabricación correspondiente.	$\pm 0,5\%$	$\pm 0,6\%$	ISO 10545-2
Ortogonalidad^b			
Máxima desviación de ortogonalidad en %, con relación a la dimensión de fabricación correspondiente.	$\pm 1,0\%$	$\pm 1,0\%$	ISO 10545-2
Planitud de la superficie			
Máxima desviación admisible de planitud en %:			
a) curvatura central con relación a la diagonal calculada con la dimensión de fabricación	$\pm 0,5\%$	$\pm 1,5\%$	ISO 10545-2
b) curvatura lateral con relación a la dimensión de fabricación correspondiente	$\pm 0,5\%$	$\pm 1,5\%$	ISO 10545-2

Tabla 23. Fragmento de la tabla de requisitos para baldosas del grupo AI (AENOR, 2007)

En la misma norma UNE-EN 14411 de 2007 (ibídem) encontramos otra tabla con información relevante para la descripción de baldosas a través de sus características titulada «Características para las diversas aplicaciones» (tabla 24). Esta tabla representa las características que resultan necesarias según el lugar de aplicación de la baldosa, además, incluye la norma que prescribe el método de ensayo para detectar y valorar cada una de las características.

De esta tabla, hemos extraído descripciones para pavimento de interior, revestimiento de interior, pavimento de exterior y revestimiento de exterior y también información sobre las características esenciales según el tipo de baldosa. Esta clasificación supone una nueva dimensión en la descripción de tipos de baldosa que hasta ahora sólo tenía en cuenta el método de moldeo y su capacidad de absorción de agua.

Características	Suelos		Paredes		Método de ensayo
	Interior	Exterior	Interior	Exterior	Referencia
Dimensiones y aspecto superficial					
Longitud y anchura	X	X	X	X	ISO 10545-2
Espesor	X	X	X	X	ISO 10545-2
Rectitud de lados	X	X	X	X	ISO 10545-2
Ortogonalidad	X	X	X	X	ISO 10545-2
Planitud de superficie (curvatura y alabeo)	X	X	X	X	ISO 10545-2
Aspecto superficial	X	X	X	X	ISO 10545-2
Propiedades físicas					
Absorción de agua	X	X	X	X	ISO 10545-3
Resistencia a la flexión	X	X	X	X	ISO 10545-4
Módulo de ruptura	X	X	X	X	ISO 10545-4
Resistencia a la abrasión profunda de baldosas no esmaltadas	X	X			ISO 10545-6
Resistencia a la abrasión superficial de baldosas esmaltadas	X	X			ISO 10545-7
Dilatación térmica lineal ^a	X	X	X	X	ISO 10545-8
Resistencia al choque térmico ^a	X	X	X	X	ISO 10545-9
Resistencia al cuarteo de las baldosas esmaltadas	X	X	X	X	ISO 10545-11
Resistencia a la helada ^b		X		X	ISO 10545-12
Coefficiente de fricción	X	X			Declarar método utilizado
Expansión por humedad ^a	X	X	X	X	ISO 10545-10
Pequeñas diferencias de color ^a	X	X	X	X	ISO 10545-16
Resistencia al impacto ^a	X	X			ISO 10545-5
Propiedades químicas					
Resistencia a las manchas					ISO 10545-14
– baldosas esmaltadas	X	X	X	X	ISO 10545-14
– baldosas no esmaltadas ^a	X	X	X	X	ISO 10545-14
Resistencia a ácidos y álcalis de baja concentración	X	X	X	X	ISO 10545-13
Resistencia a ácidos y álcalis de alta concentración ^a	X	X	X	X	ISO 10545-13
Resistencia a los productos domésticos de limpieza y aditivos para agua de piscinas	X	X	X	X	ISO 10545-13
Emisión de cadmio y plomo de baldosas esmaltadas ^a	X	X	X	X	ISO 10545-15
^a Método disponible, aunque esta norma no especifica valores.					
^b Para baldosas que se destinan a ser utilizadas en lugares donde pueden estar sometidas a condiciones de hielo.					

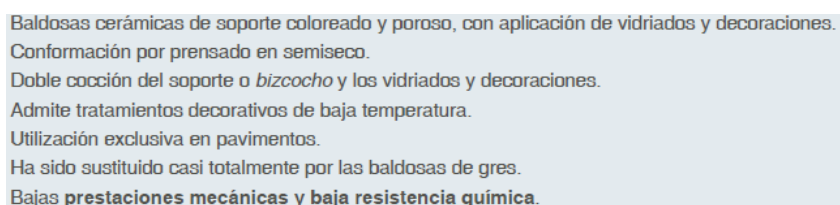
Tabla 24. Tabla de la norma UNE-EN 14411 sobre características según la aplicación de la baldosa (AENOR, 2007:13)

5.3.2 Clasificación comercial de las baldosas cerámicas de PROALSO

La clasificación comercial de las baldosas cerámicas de PROALSO (PROALSO, 2011) se ha utilizando para describir los tipos de baldosa presentes en la clasificación y probar si el propio razonador reclasificaba estas baldosas en los correspondientes grupos de la norma UNE-EN 14411 (AENOR, 2007) tal y como figuran en la misma tabla. Al disponer de la equivalencia posibles entre las baldosas descritas y las de la norma UNE 14411, nos ha permitido probar la adecuación de nuestras descripciones y evaluar la consistencia de la ontología en este aspecto.

Para ello hemos recogido los datos de cada baldosa descrita en el documento y los hemos incluido en una ficha de baldosas. Los datos de esta ficha se introducirán después en la ontología, obviando la referencia al grupo o grupos según la norma, para ver si las baldosas se reagrupan bajo la clase indicada.

Es necesario remarcar que los resultados de extracción de descripciones de las baldosas de este tipo no tienen el nivel de detalle que tienen las clasificaciones de las normas UNE de cerámica, que describimos a continuación. Hemos notado que las descripciones presentes en los textos se basan en indicar la diferencia de un tipo de baldosa con respecto a otro, con lo que se omiten características esenciales en muchos de los casos y se mencionan únicamente las distintivas. Por ejemplo, la descripción de ‘pavimento de bicocción’ en el manual es la siguiente (figura 82):



Baldosas cerámicas de soporte coloreado y poroso, con aplicación de vidriados y decoraciones.
Conformación por prensado en semiseco.
Doble cocción del soporte o *bizcocho* y los vidriados y decoraciones.
Admite tratamientos decorativos de baja temperatura.
Utilización exclusiva en pavimentos.
Ha sido sustituido casi totalmente por las baldosas de gres.
Bajas prestaciones mecánicas y baja resistencia química.

Figura 79. Descripción de ‘pavimento de bicocción’

En la última línea de la descripción se indica que el pavimento de bicocción tiene bajas prestaciones mecánicas y baja resistencia química, pero no especifica con respecto a qué se considera baja. En el caso de ‘resistencia química’ los valores posibles serían GA, GB o GC, siendo GA el valor más resistente y GC el que indica menor resistencia. Podemos interpretar que baja resistencia química hace referencia al valor GC pero no podemos afirmarlo sin consultar con un experto. Al mismo tiempo, se indica un proceso de moldeo que es más específico de lo que generalmente suele indicarse en catálogos o normas de estandarización, que distinguen entre extrusión y prensado en seco. Tal y como se redacta parece excluir a los pavimentos con doble cocción cuyo proceso de moldeo sea el prensado en seco, aspecto que contrastado con el resto de recursos textuales, no queda probado.

5.3.3 Resultados y observaciones

Este procedimiento ha dado como resultado un conjunto de fichas descriptivas de baldosas a través de características presentes en los recursos utilizados (anexo 2). Además, al extraer la información que permite describir los tipos de baldosa según la norma, hemos identificado una relación entre la baldosa y dos procesos relacionados con la fabricación de la baldosa que resultan determinantes en su clasificación. Esta relación (relación baldosa cerámica – proceso de fabricación) se ha incluido en la ontología de manera que la descripción de los tipos de baldosa sea consistente y veraz.

5.4 Consulta a experto y validación

Para corroborar la adecuación de las características extraídas y de la información recogida y resolver los conflictos que nos hemos ido encontrando durante la fase de recogida de información hemos recurrido a un experto del campo de la cerámica. El experto es Isaac Nebot-Díaz, Doctor en Química por la Universidad Jaume I, que ejerce de jefe de estudios de la Escuela Superior de Cerámica de la Alcora desde hace más de seis años. Entre sus numerosas aportaciones al campo de la cerámica destaca su tesis doctoral sobre nuevas rutas de síntesis para pigmentos cerámicos (Nebot-Díaz, 2001) o el libro *Nuevas tecnologías para el sector*

cerámico (Nebot-Díaz et al., 2000). Además, cuenta con una amplia experiencia en el sector privado, habiendo trabajado como en el departamento de I+D+I de empresas como VENIS S.A. o Abissal Invest, SL.

El Dr. Nebot-Díaz nos recibió en la Escuela Superior para realizar una serie de entrevistas que permitieran resolver las dudas surgidas durante la fase de recogida de datos y validar los datos recogidos. La primera entrevista tuvo lugar el día 26 de junio de 2012. A continuación, explicamos cómo se llevo a cabo la entrevista y qué resultados obtuvimos.

En primer lugar, se explicó al Dr. Nebot-Díaz el objetivo de nuestra investigación y el trabajo realizado. Asimismo, se explicó cuál era el objetivo de la entrevista y la metodología propuesta para la resolución de dudas y la validación. La primera parte, consistió en la validación general de los grandes tipos de conceptos incluidos en el trabajo: las características, los procesos relacionados con las características, los instrumentos relacionados con las características, las baldosas que se describirían y otros aspectos importantes para enmarcar el trabajo.

A continuación, se procedió a la consulta de dudas puntuales sobre equivalencias de denominación, validación de valores para una característica o validación de la relevancia de las características para según qué tipo de baldosa. El experto resolvió pacientemente nuestras consultas y además, nos ejemplificó muchas de las observaciones realizadas con la visita al laboratorio de calidad de la baldosa del centro (donde se mesuran la mayor parte de las características de la baldosa). En este laboratorio y con la instrucción del experto, pudimos apreciar el funcionamiento de muchos de los instrumentos referidos a las características, características que podrían observarse en baldosas presentes en el laboratorio, procedimientos vinculados a las características, etc. Todo esto, supuso una gran fuente de conocimiento añadido que sirvió para completar la información contenida en nuestra ontología.

Finalmente, se acordó entregar al Dr. Nebot-Díaz los documentos de trabajo elaborados durante toda esta fase de extracción (tabla de característica, fichas de baldosas, listados). La finalidad de esta entrega era que el Dr. pudiera analizar el contenido de los documentos y proceder a la validación de datos. Igualmente, se acordó que en una segunda entrevista fijada para el 12 de julio, el experto expondría los resultados de su análisis y validación.

En referencia a los resultados de la primera entrevista, en los siguientes párrafos presentamos las dudas consultadas más relevantes para la creación de la ontología y las observaciones del experto.

Uso de valores “precisión” o “natural” que aparecen en las Normas UNE-EN ISO.
El problema que habíamos detectado es que al indicar los valores mínimos de las características que describen un determinado tipo de baldosa, en las tablas de clasificación se indicaban dos tipos de valor “precisión” y “natural” (tabla 25).

Propiedades químicas	Precisión	Natural	Ensayo
Resistencia a las manchas			
a) baldosas esmaltadas	Mínimo clase 3	Mínimo clase 3	ISO 10545-14
b) baldosas no esmaltadas	Método disponible	Método disponible	ISO 10545-14
Resistencia a productos químicos			
Resistencia a ácidos y álcalis de baja concentración			
a) baldosas esmaltadas	Según la clase indicada por el fabricante	Según la clase indicada por el fabricante	ISO 10545-13
b) baldosas no esmaltadas ^g			
Resistencia a ácidos y álcalis de alta concentración ^e	Método disponible	Método disponible	ISO 10545-13
Resistencia a productos domésticos y sales para piscinas			
a) baldosas esmaltadas	Mínimo GB	Mínimo GB	ISO 10545-13
b) baldosas no esmaltadas ^g	Mínimo UB	Mínimo UB	
Emisión de plomo y cadmio^f	Método disponible	Método disponible	ISO 10545-15

Tabla 25. Extracto de la tabla de requisitos para baldosas extruidas, Grupo AI, E₃ según norma UNE-EN

14411

Cuando el experto fue preguntado acerca de esto, explicó que depende del aparato de medida utilizado. Las medidas aludidas bajo el epígrafe “precisión” hacen referencia a medidas tomadas normalmente en laboratorios, en un entorno estable y seguro, donde se puede hacer uso de aparatos de alta precisión sin que estos se dañen o sufran variaciones de temperatura que afecten a su rendimiento. En cambio, en otros casos, las medidas se realizan en la misma planta de producción cerámica, con unas condiciones ambientales que no permiten el uso de estos aparatos sino de otros más resistentes pero, a menudo, menos precisos. La recomendación del experto es centrarse en los valores indicados en la columna de “precisión” ya que a día de hoy cualquier producto cerámico comercial es testado en laboratorio para certificar los mínimos de la norma UNE-EN ISO.

Diferencia entre algunos subgrupos de baldosas establecidos en la norma UNE 14411. Los criterios de clasificación de los subgrupos AII_{a-1} y AII_{a-2} de baldosas no parecían diferir el uno del otro si comparáramos los valores de la columna “natural” de las tablas de clasificación de la norma UNE-EN 14411 (AENOR, 2007). Las observaciones del experto respecto al punto anterior nos confirman la preferencia de los valores indicados en la columna “precisión” como los valores a tener en cuenta.

Validación de instrumentos recogidos por ser demasiado generales. El experto revisó la lista de instrumentos vinculados a la medida de las características y validó la información aportando, además, equivalentes en español o términos más específicos que los aportados en la lista. Por ejemplo, el experto explica “el ‘plucómetro’ es el equipo utilizado para medir el

espesor, la ortogonalidad, planitud de la superficie” y contiene el ‘micrómetro’ (término incluido en nuestra lista de instrumentos de medida de las características).

Valores de características no expresados en algunos catálogos y referenciados como “Método disponible”. Durante el análisis se han dado casos donde, en las tablas de descripción de las baldosas en los catálogos o las normas, el valor de una característica no se especifica sino que aparece una referencia que indica que el método está disponible. Por ejemplo, en la tabla 26 se muestra un fragmento de la tabla que describe las baldosas extruidas del grupo AII_{a-2} en la norma UNE-EN 14411.

Dilatación por humedad, en mm/m*	Método disponible	Método disponible	ISO 10545-10
Pequeñas diferencias de color*	Método disponible	Método disponible	ISO 10545-16
Resistencia al impacto*	Método disponible	Método disponible	ISO 10545-5

Tabla 26. Fragmento de la tabla de descripción de baldosas extruidas del grupo AII_{a-2}

El experto explicó que esto se debe a temas de coste por los ensayos requeridos para certificar el valor de estas características. Hay que tener en cuenta que los ensayos que certifican los valores de un tipo de baldosa se deben realizar por una entidad normalizadora y que cada ensayo para cada tipo de baldosa producida en fábrica implica un coste. Es por eso que, en principio, estas características al no resultar delimitadoras en el tipo de baldosa descrita no requieren el ensayo por parte de una entidad normalizadora.

La clasificación de baldosas. Nos preocupaba que la clasificación de tipos de baldosa quedara muy limitada ya que sólo hemos encontrado la clasificación de tipos de baldosa según grupos de la norma UNE-EN 14411. El experto señaló que esa era la clasificación normativa y la utilizada en la industria cerámica. A la pregunta por el uso de denominaciones de tipos de baldosa que habíamos encontrado en el corpus, concretamente, dos tipos de baldosa para los que se utiliza su denominación en inglés (‘quarry tile’ y ‘split tile’), el experto señaló que esos dos tipos de baldosa están en desuso y que independientemente, lo que valía de cara a la comercialización era el tipo según la normativa. En referencia a los tipos de baldosa encontrados en los catálogos de cerámica que suelen ir vinculados a un determinado grupo según la norma (por ejemplo, baldosa extruida, baldosa de gres, baldosa de gres porcelánico), el experto explicó que se trata de otras denominaciones que hacen referencia a los grupos según la norma, por ejemplo, señaló “está el gres, la monoporosa, el porcelánico y hay que ver cada una de ellas a qué grupo corresponde según sus cualidades de absorción y moldeo [...], por ejemplo, el grupo I está hecho por extrusión, hace referencia a *baldosas extruidas*“. La conclusión a que llegamos es que existen equivalencias entre los grupos de la norma y los tipos de baldosa encontrados en los catálogos.

La diferenciación entre los tres grupos (I, II y III). En la clasificación que establece la norma UNE-EN 14411 sobre tipos de baldosa, el grupo I corresponde a baldosas elaboradas mediante el método de extrusión, el grupo II a baldosas elaboradas por el método de prensado en seco y el grupo III hace referencia a baldosas elaboradas por “otros métodos”. Nos preocupaba la diferencia de exactitud entre los grupos I y II y el grupo III. El experto, preguntado por este asunto, nos explicó que el grupo III, “denominado también *otros conformados*”, es efectivamente un cajón de sastre, hay que tener en cuenta prácticamente todas las baldosas fabricadas con fines comerciales se incluyen en los grupos I y II. Además, el

experto nos advierte de otra tipología que se basa en el tipo de pasta utilizada, pasta blanca y pasta roja (según la arcilla utilizada), pudiendo ser gres de pasta blanca, gres de pasta roja, monoporosa de pasta blanca, monoporosa de pasta roja, etc. Esto conduce a un amplio debate sobre la prevalencia de la pasta blanca (presente en áreas de Italia y China) sobre la pasta roja (propia de la cuenca del valle de Alcora) por cuestiones de marketing y nos conduce a la siguiente cuestión.

La composición mineralógica de la baldosa. Conscientes de la importancia de la composición mineralógica de la baldosa para su diferenciación con respecto a otro tipo de pavimentos o revestimientos y entre tipos de baldosa, transmitimos al experto nuestra preocupación para encontrar valores concretos que hagan referencia a esta característica. El experto nos explica que todas las baldosas se componen de arcillas y la arcilla se compone de caolín, feldespato y cuarzo; según la proporción serán arcillas caoliníticas (alto contenido en caolín), arenas cuarcíferas (alto contenido en cuarzo). En este aspecto, el experto señala que la diferencia importante radica en las cantidades de óxido de hierro presentes en las arcillas, que dan lugar a dos grandes tipos de baldosa, en base a su composición: las baldosas de pasta blanca y las baldosas de pasta roja. La diferencia entre pasta blanca y pasta roja depende del óxido de hierro que tengan las arcillas, a más óxido de hierro más roja es la pasta. Por tanto, concluimos que nuestra ontología debe incluir la característica ‘composición’ de una baldosa y que sus valores serán ‘pasta blanca’ y ‘pasta roja’. Además, se incluirá una descripción de lo arriba indicado para recoger la composición de las arcillas que conforman la pasta roja y pasta blanca y la diferencia entre estas pastas.

Valores de la característica ‘acabado’. Se preguntó al experto por la distinción entre los valores de ‘acabado’ que habíamos encontrado y parecían pertenecer a dos tipologías; por una parte, esmaltado, no esmaltado ;por otra, brillo, satinado, mate, etc. El experto confirma que existe la distinción y señala que habría que hacer una primera diferenciación entre dos tipos de acabado, esmaltado y no esmaltado, y a continuación entre los tipos de esmaltado: brillo, satinado, mate, etc.

Equivalencia de denominaciones. Se preguntó al experto por distintas denominaciones encontradas en el análisis que parecían hacer referencia al mismo concepto. El experto confirmó, desmintió o aclaró las distintas equivalencias encontradas. Por ejemplo, *resistencia a la flexión y fuerza de rotura* como equivalentes, *carga de rotura y módulo de ruptura* no son equivalentes, *resistencia a los productos domésticos de limpieza* suele ser la forma corta de *resistencia a los productos domésticos de limpieza y aditivos para piscina*.

Defectos vinculados a la calidad superficial. El experto valida los defectos encontrados y explica que existen una variedad de denominaciones para hacer referencia a los mismos defectos, algunos son préstamos del inglés (*pinhole* como sinónimo de *pinchado*) y otros se deben a las características sociolingüísticas de la zona donde conviven el español y el catalán (por ejemplo, *punxat*, del catalán, se utiliza en castellano para hacer referencia al defecto *pinchado*). Resultaría muy interesante explorar estos préstamos pero las circunstancias nos obligan a establecer límites en este trabajo.

Color. Transmitimos al experto la dificultad para encontrar una clasificación de colores de la baldosa consistente y objetiva. El experto nos explica que en el laboratorio la clasificación de colores se basa en el espacio de color CieLab pero que luego, por razones de marketing, un color con las mismas coordenadas CieLab, puede tener numerosas denominaciones, por

ejemplo: marfil, arena del Sáhara, moca marengo, etc. Se trata de una cuestión de marketing que da lugar a infinitos valores para color. Estas observaciones resultaron muy aclaratorias y nos comprometimos a presentar una propuesta que tuviera en cuenta estos dos aspectos de color: el espacio de color CieLab (utilizado en laboratorio) y los tipos de color para el producto acabado (utilizados en entornos comerciales y por los prescriptores cerámicos).

La segunda reunión tuvo lugar el 12 de julio en la Facultad de Ciencias Humanas y Sociales de la Universidad Jaume I. Durante esta reunión el experto expuso los resultados de la lectura y análisis de los materiales de trabajo que le habíamos presentado en la reunión anterior. La valoración del experto fue muy positiva, probando la validez y realizando las siguientes observaciones:

- En la representación, debe mantenerse una diferenciación entre los procesos ‘moldeo’ y ‘cocción’ y las características de la baldosa. Estos procesos, aunque son anteriores a la finalización de la baldosa cerámica, resultan relevantes en la clasificación y descripción de las baldosas y mantienen relaciones con características de éstas.
- Revisada toda la lista de características, y los valores de éstas, el experto concluye que son correctas pero faltaría añadir la característica ‘resistencia al rayado’, cuyos valores se miden según la escala de dureza Mohs (de 1 a 10, siendo 10 la dureza del diamante)

Además de la validación, aprovechamos para transmitir al experto las decisiones que habíamos tomado a raíz de la reunión anterior y hacerle una propuesta para representar los valores ‘color’. La formalización de color tendría en cuenta dos posibilidades: expresar los valores de color a través de denominaciones utilizadas en el plano comercial y agregar la posibilidad de enlazar estas denominaciones con las coordenadas en el espacio CieLab. El experto valora la propuesta y considera que así se mantiene la coherencia dentro de lo posible aún sabiendo que las denominaciones para los valores de color en cerámica son una cuestión de marketing con un carácter altamente subjetivo y comercial.

Con esta reunión concluyó la etapa de validación de datos y se modificaron los documentos de trabajo para llegar a la versión final que presentamos en los anexos 1-7. Es necesario remarcar que la colaboración del experto supuso una ayuda inestimable, capaz de resolver dudas en el momento y advertir sobre otros posibles conflictos futuros, su colaboración ha contribuido enormemente a la calidad del contenido formalizado. Para finalizar el experto añadió valor a nuestro trabajo recomendándonos una futura vía alternativa de explotación de la ontología orientada a los prescriptores de la cerámica: “arquitectos, almacenes de distribución, despachos de decoración, es decir, los últimos intermediarios antes de poner la baldosa o adquirirla” que podrían estar interesados en este tipo de trabajo.

5.5 Resultados

Los resultados de esta fase de extracción de características y validación consisten en una tabla de datos Excel validada donde se recogen las 41 características del concepto ‘baldosa cerámica’ y sus correspondientes contextos, definiciones e información relevante para la clasificación. Además, obtenemos información sobre la naturaleza de las características que resultará de gran interés para describir los elementos que forman parte de la descripción de una característica y las relaciones y tipos de datos que podemos encontrar en su descripción.

En la descripción de las características participan 16 procesos, 47 parámetros, 83 instrumentos y 8 unidades de medida. Hemos recogido 143 contextos en total del corpus TxtCeram, las normas o los catálogos que describen las características o procesos o parámetros de las características.

Además, hemos recogido la descripción de 41 tipos de baldosa que se implementarán en la ontología y hemos identificado una nueva relación, la relación *baldosa cerámica – proceso de fabricación*, que resulta relevante para describir los tipos de baldosas. En este sentido, aunque la formalización de una relación y una característica resulta similar como se observará más adelante, hemos optado por representar el binomio baldosa cerámica-proceso de fabricación como una relación y no una característica para dejar abierta la posibilidad de ampliar la rama de procesos de la baldosa y conectarla con la baldosa a través de relaciones.

Respecto a otras denominaciones para características, hemos recogido 21 denominaciones sinónimas para 20 características. De manera simbólica, se han recogido los equivalentes en inglés para las 41 características de la baldosa cerámica y las 6 clases que las agrupan.

6 Formalización de las características en la ontología

6.1 Formalización con Protégé en OWL 2

Las ontologías, vistas como un sistema que permite la conceptualización y por tanto, la representación de conocimiento, se basan, generalmente, en sintaxis y lenguajes formales que permiten expresar dicho conocimiento. Una vez identificados y definidos los elementos que formarán parte de la descripción de una característica y las relaciones que se mantienen entre éstas y otros conceptos, hemos formalizado este conocimiento haciendo uso de un editor de ontologías y hemos ‘traducido’ la naturaleza de cada uno de los elementos que forman parte de nuestra descripción, hasta ahora expresados en lenguaje natural, a un lenguaje más controlado y formal.

Para hacer esto, podemos utilizar distintos sistemas que a su vez hacen uso de distintos lenguajes: la lógica de descripciones, la lógica de primer orden, etc. En nuestro trabajo hemos optado por utilizar el lenguaje OWL 2 para describir las características.

Mediante el lenguaje OWL 2, y desde la herramienta Protégé 4, vamos a engranar todos los elementos y facetas de una característica. Este paso es delicado ya que las funcionalidades de la herramienta y de OWL 2 han tenido que adaptarse a nuestras necesidades descriptivas. El lenguaje OWL 2 no dispone de equivalentes directos para cubrir nuestros propósitos de descripción, por lo que no siempre hay una forma directa de realizar el traspaso entre lo que hemos descrito en lenguaje natural y el lenguaje formal OWL 2. Por eso, para llevar a cabo esta transferencia de información hemos tenido que determinar de entre las distintas posibilidades de formalización que ofrece OWL 2 cuáles resultan más cercanas y acertadas para nuestros objetivos de formalización.

Para llevar a cabo la representación formal en OWL 2 de una forma más cómoda y sencilla, hemos utilizado la herramienta de edición Protégé 4 (véase II.10.1.5). Esta herramienta nos permite utilizar la sintaxis del lenguaje OWL 2 de acuerdo con distintos perfiles (lite, DL y full) que introducen mayor o menor expresividad a la ontología. En este caso para poder alcanzar un estándar en la medida de lo posible vamos a utilizar las expresiones propias de OWL 2 DL.

Para elaborar las descripciones de las clases de la ontología y de las propiedades, se ha utilizado la *Manchester OWL 2 Syntax*, con lo que la complejidad a la hora de describir las facetas de nuestras características se ha visto aligerada. Para asegurar la consistencia de la ontología y realizar las consultas hemos utilizado el razonador *HERMIT*, que además nos ha ayudado a inferir información a partir de lo representado de manera que se evita declarar información redundante.

6.2 Formalización de clases en la ontología y aspectos generales

Para favorecer el seguimiento de este capítulo y los que siguen, recomendamos consultar la ontología y el anexo 11 sobre instalación del software y acceso a la ontología.

La ontología está enfocada a la descripción de las características de la baldosa cerámica con lo que éstas suponen el centro de atención de nuestro trabajo. No obstante, tal y como

hemos visto anteriormente, en la descripción de la característica participan otro tipo de conceptos que están presentes en la ontología, por ejemplo: entidades que representan tipos de baldosa, actividades relacionadas con las características o relaciones conceptuales que se establecen entre características o entre características y otros conceptos. Por eso, al formalizar esta información en la ontología, hemos diferenciado entre estos elementos que representan conceptos y también, dentro de un mismo concepto, especialmente en características, entre sus distintas facetas.

Cada uno de los conceptos formalizados en la ontología cumple con un rol o función dentro de ésta. Su formalización variará según el rol y dependerá de si el concepto o la información a formalizar se puede expresar más adecuadamente a través del elemento *class*, *object property*, *datatype property*, *individual*, *annotation*²⁷; o mediante una expresión que aúne y combine distintos elementos (por ejemplo, una expresión que atribuye una propiedad a una clase).

Para entender la idea general en la que se basa la ontología podemos partir de lo que observamos desde el editor de ontologías Protégé 4, que es la aplicación que hemos utilizado para formalizar la información en OWL 2 e implementarla. En el editor, encontramos distintas pestañas, y entre ellas, cuatro que se corresponden con cuatro tipos de entidades OWL 2 (OWL 2 entity). De acuerdo con la versión inglesa de la aplicación estas pestañas se denominan: *Classes*, *Object Properties*, *Datatype Properties* e *Individuals*. Así, por ejemplo, la pestaña de *Classes* contiene las clases de la ontología y su descripción, o la pantalla *Object properties* contiene los elementos de tipo *object property* y datos sobre su configuración (dominio, rango, etc.).

Si nos situamos en la pestaña de las clases, una perspectiva general de la ontología (figura 85) nos muestra que el primer nivel jerárquico de clases se compone de características (donde se incluirán las características del producto acabado, los parámetros que describen las características, la clase que representa a ‘característica intrínseca’ y la clase que representa a ‘característica extrínseca’), entidades (donde se incluirán los tipos de baldosa, los instrumentos de medida de las características y las unidades de medida) y actividades (donde se incluirán procesos relacionados con las características y con las baldosas). Además de las tres grandes clases de conceptos (entidad, actividad y característica), nuestra ontología incluye la clase ‘cerámica’ que representa el ámbito y agrupará todos los conceptos relacionados con este ámbito.

Es decir, que la formalización de las características en la ontología parte de la creación de una clase para cada tipo de concepto involucrado en su descripción y para el ámbito de especialidad.

²⁷ En adelante y siempre que esto no produzca ambigüedad, nos referiremos al elemento *class* como clase, a los elementos *object property* y *datatype property* como propiedad, al elemento *individual* como instancia y al elemento *annotation* como anotación.

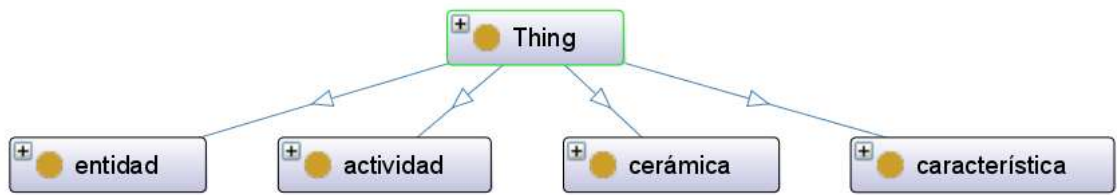


Figura 80. Grafo del primer nivel de la jerarquía de clases de la ontología

En el extracto de grafo generado a partir la ontología, observamos el primer nivel de la jerarquía de clases. Las relaciones no se han representado en forma de clase en la ontología sino sólo únicamente en su rol descriptor en forma de propiedad. Esto se debe a que la información que queríamos formalizar en la ontología se ha podido formalizar a través de propiedades sin necesidad de representar las relaciones en forma de clase (véase §6.9).

Los conceptos subordinados a ‘característica’, ‘entidad’ y ‘actividad’ se han formalizado a través de subclases de estas clases y en el caso de tener una relación con el ámbito de la cerámica, también se han formalizado como subclases de cerámica. Por ejemplo, (figura 74).

Un segundo nivel de la jerarquía en el que se desarrollan las subclases de ‘cerámica’ puede verse en la figura 74. En el grafo representado en esta figura, aparece la relación de subordinación entre la clase ‘cerámica’ y las subclases de ‘entidad’, ‘característica’ y ‘actividad’ de la ontología en forma de flechas oscuras, estas mismas clases pertenecen al mismo tiempo a otras clases como se visualiza en el grafo a través de las flechas más claras. En el anexo 8 se puede visualizar un grafo con las superclases de la ontología.

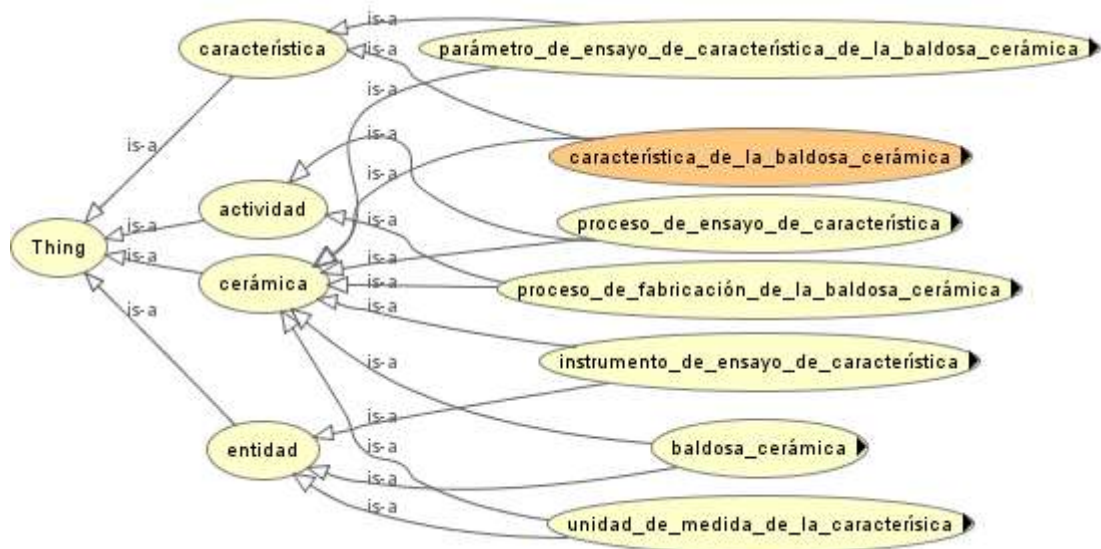


Figura 81. Grafo de la jerarquía de segundo nivel de la clase ‘cerámica’

En el caso que dos o más clases hermanas no pudieran compartir miembros entre ellas, las clases disjuntas se han declarado en la descripción de las clases implicadas. Por ejemplo, la clase ‘actividad’ se ha formalizado como disjunta de ‘entidad’ y ‘característica’. La formalización se realiza desde la ventana de descripción de la clase en el apartado *Disjoint classes* (figura 82).



Figura 82. Formalización de clases disjuntas

Además de las clases de la ontología, el rol descriptor de las características y las relaciones conceptuales presentes en la ontología se han formalizado a través de *object property* o *datatype property*.

Las *object properties* que componen la ontología se visualizan desde la pestaña *Object Properties* de la Protégé. Una *object property* tiene la función de vincular mediante una etiqueta descriptiva a dos clases de la ontología, con lo que describe a una de ellas atribuyendo como valor de la etiqueta a la otra. Para ello, hemos asignado dominios y rangos a cada *object property* tal y como se visualiza desde el apartado *description* en Protégé. Por ejemplo, la *object property* ‘tiene_acabado’ tiene como dominio la clase ‘baldosa cerámica’, es decir, describe a baldosas y como rango la clase ‘acabado’, es decir, el valor de la propiedad tendrá que ser algún miembro de la clase ‘acabado’.

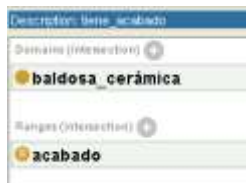


Figura 83. Ejemplo de jerarquía desde la pestaña *object property*

En el primer nivel de la jerarquía de *object properties* hemos distinguido tres grandes grupos: características de la baldosa cerámica, características de la característica de la baldosa cerámica y relaciones (figura 84). El grupo de ‘características de la característica de la baldosa’ contiene *object property* que representan las características que hemos utilizado para describir las características de la baldosa. El grupo de ‘características de la baldosa’ incluye *object property* que representan el rol descriptor de algunas características de la baldosa. Finalmente, el grupo de ‘relaciones’ incluye *object property* que representan la relación de meronimia entre una característica y otra, la relación entre una característica y el proceso de ensayo utilizado para determinar su valor y la relación entre una baldosa cerámica y el proceso de fabricación utilizado.

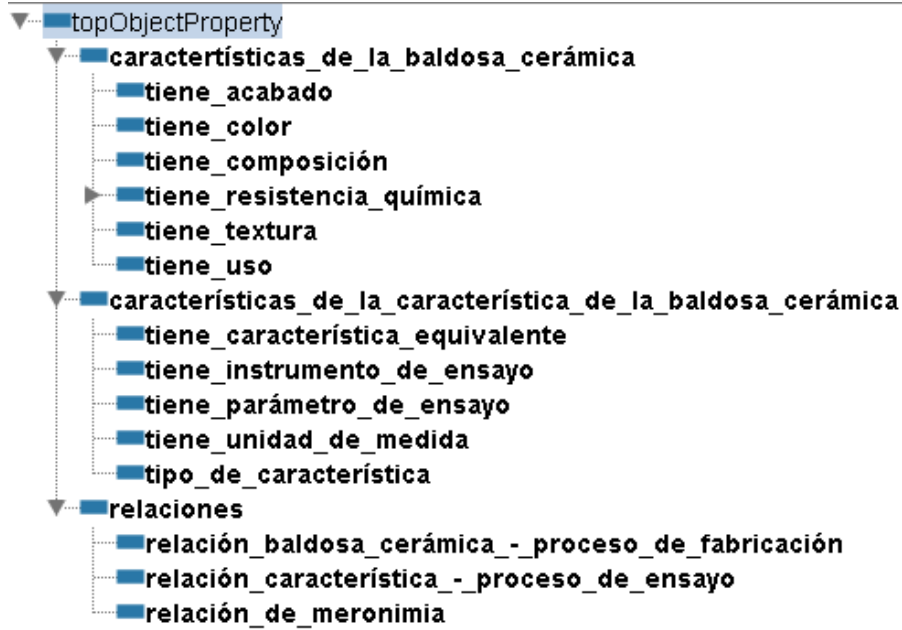


Figura 84. Jerarquía de *object property* en la ontología

En la pestaña *Data properties* del programa, podemos ver las características cuyo rol descriptor se ha formalizado a través de *datatype property*. Estas características son aquellas cuyos valores se refieren a números o valores booleanos y no aparecen formalizados como clases o instancias. Al formalizar estas características a través de *datatype property* hemos representado que la característica describe a un determinado tipo de concepto (representado en forma de clase) y que toma un determinado tipo de valor o un valor concreto. Por ejemplo, en la figura 85 podemos ver la descripción de ‘tiene aspecto superficial’ que describe a baldosas y cuyos valores son números enteros.



Figura 85. Descripción de la *datatype property* para ‘aspecto superficial’

En nuestra ontología, todos los elementos de tipo *datatype property* aluden a una característica y se han agrupado bajo la propiedad ‘característica de la baldosa’, como se puede observar en la siguiente imagen (figura 86).

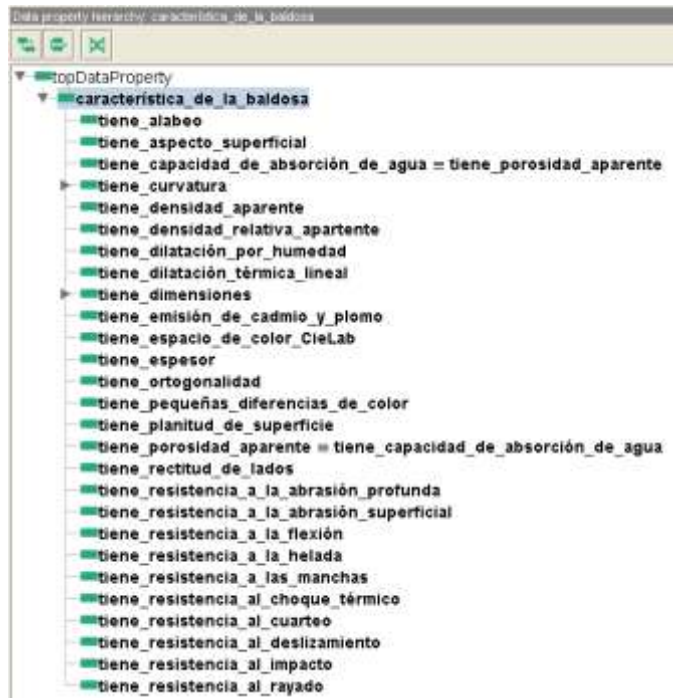


Figura 86. Datatype properties que representan características

Además de los elementos de tipo *class* y *property*, en la ontología conceptos formalizados en forma de *individual* (o instancia). Las instancias no están jerarquizadas entre ellas sino que están relacionadas con alguna de las clases de la ontología que permite agruparlas y clasificarlas. Cada una de las instancias de la ontología representa un valor concreto de una característica y se agrupa bajo la clase de característica.



Figura 87. Ejemplo de instancias de la subclase de acabado 'esmaltado' y descripción de la instancia 'alto brillo' como un tipo de 'esmaltado', visible desde la pestaña

En la ontología también hemos utilizado *annotation properties* para aquellos casos en los que sea necesario incluir información en forma de anotación sin que por ello ésta deba ser considerada por los razonadores para estructurar la ontología. Consiste en información en lenguaje natural sobre las clases, instancias o propiedades de la ontología como, por ejemplo, un contexto de la característica. En el editor Protégé, podemos visualizar la información formalizada a través de anotaciones en la correspondiente subventana *Annotations*. Por ejemplo, en el caso de las clases, lo encontramos en la subpestaña *Annotations* de la pestaña *Classes*.

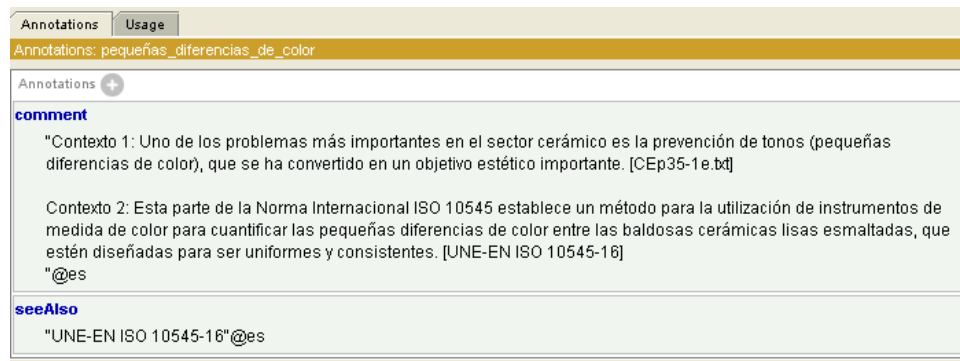


Figura 88. Ejemplo de *annotation properties* incluidas en la descripción de la clase *pequeñas diferencias de color*

En la figura 88, podemos observar dos anotaciones. Una a través del subelemento de anotación *comment* y otra a través del subelemento de anotación *seeAlso*. A través de *comment* hemos representado los contextos de una característica en lenguaje natural y a través de *seeAlso* la fuente de la característica. Más adelante, profundizamos en los detalles de estas formalizaciones.

Finalmente, las descripciones relativas a las características y las relaciones de las clases de características de la baldosa y de baldosa cerámica se han formalizado mediante expresiones que combinan el uso de propiedades con primitivas y restricciones a través de los distintos apartados incluidos en la sub-ventana de descripción de las clases. Estas expresiones nos han servido para traducir las descripciones en lenguaje natural sobre una clase de concepto a descripciones formales en OWL 2. A modo de metáfora, podemos entender que estas expresiones actúan como el cemento que une los ladrillos, las ventanas y las puertas de una casa. Así, se van estableciendo los límites de la ontología, atribuyendo una característica a una clase, restringiendo los miembros de la clase, indicando que existe una equivalencia entre clases, etc.

La descripción de cada clase de la ontología varía según el tipo de clase que se trate pero, a modo de introducción, exponemos la formalización de la clase 'espesor'.

The screenshot shows a window titled 'Description: espesor'. It contains three main sections:

- Equivalent classes:** A single entry: **● tiene parámetro de ensayo *some* grosor_medio_de_los_lados**
- Superclasses:** Four entries:
 - medida**
 - tiene instrumento de ensayo *some* micrómetro_con_palpadores**
 - tiene parámetro de ensayo *only* grosor_medio_de_los_lados**
 - tiene unidad de medida *some* milímetro**
- Inherited anonymous classes:** Six entries:
 - tiene parámetro de ensayo *only* parámetro_de_ensayo_de_característica_de_la_baldosa_cerámica**
 - tiene parámetro de ensayo *some* parámetro_de_ensayo_de_característica_de_la_baldosa_cerámica**
 - apariencia**
 - or* característica_física
 - or* característica_química
 - or* composición
 - or* medida
 - or* uso
 - tiene parámetro de ensayo *only* (grosor_medio_de_los_lados**
 - or* medida_de_lados_cortos
 - or* medida_de_lados_largos
 - or* porcentaje_de_desviación)
 - tipo_de_característica *some* característica_intrínseca**
 - tiene parámetro de ensayo *some* (grosor_medio_de_los_lados**
 - or* medida_de_lados_cortos
 - or* medida_de_lados_largos
 - or* porcentaje_de_desviación)

Figura 89. Descripción de la clase 'espesor'

La característica distintiva de una clase, se formaliza combinando los elementos *some* y *only* en los apartados *Equivalent classes* y *Superclasses* respectivamente. En la figura 89, la información que se formaliza en el apartado de *Equivalent classes* hace referencia a la característica distintiva de 'espesor' respecto a sus hermanas de clase, en este caso, el parámetro de ensayo. Esta declaración se complementa con la expresión "tiene parámetro de ensayo *only* grosor medio de los lados" para indicar que la característica distintiva exige además el valor 'grosor medio de los lados' para la clase 'espesor'. Es decir, lo que distingue a 'espesor' de sus hermanas de clase es tener un parámetro de ensayo y que éste tenga el valor 'grosor medio de los lados'.

Desde *Superclasses* hemos formalizado también la información que resulta esencial como miembro de una clase. En este apartado, se declara que la clase a la que pertenece 'espesor' es 'medida', atribuyéndola como superclase. Además, hemos indicado que la clase tiene una unidad de medida que es 'milímetro' y que el instrumento de ensayo utilizado para medir el espesor es 'micrómetro con palpadores'. Esta información se formaliza a través de las expresiones 'tiene unidad de medida *some* milímetro' y 'tiene instrumento de ensayo *some* micrómetro con palpadores', el elemento *some* indica que en caso de tener ese atributo uno de

los valores debería ser el explicitado. Si alguna otra característica en la ontología cumpliera con todos estos requisitos, esa característica se consideraría equivalente a la característica ‘espesor’.

Además, en la figura 89, se visualizan otras características de ‘espesor’ que ha heredado como miembro de la clase ‘medida’ y han sido inferidas por el razonador.

A continuación, vamos a presentar con detalle cómo hemos formalizado toda la información presente en la ontología, porqué y mediante qué elementos y mecanismos de los arriba descritos.

6.3 Formalización de característica: nombre y valor

Además de representar las características como clase conceptual y estructurarlas en forma de clase y subclase, hemos representado su papel como elemento descriptor haciendo uso del elemento *property*. En OWL 2²⁸, el elemento que sirve para describir conceptos se denomina *property* y se utiliza para establecer un enlace con carga descriptiva entre un concepto de la ontología y otro o entre un concepto de la ontología y un tipo de dato. Decimos que se trata de un enlace con carga descriptiva porque el enlace en sí mismo recibirá un nombre que identifica el vínculo que se mantiene entre dos conceptos. Este elemento se utiliza para describir aspectos de un concepto, los cuales desde un punto de vista terminológico, pueden ser tanto relaciones del concepto como características de éste.

Para que la OWL 2 *property* cumpla su papel descriptor en la ontología, a ésta se le asigna un dominio y un rango. En el caso de formalizar una característica, el dominio de la *property* es el concepto o clase conceptual que es descrito por la característica y el rango indica el valor, conjunto de valores o tipo de valor de la característica que se aplica a ese dominio. En la ontología, la característica se formaliza, a este nivel, añadiendo la designación de ésta en forma de elemento *property*. Esta designación suele ir acompañada de una aclaración sobre el tipo de información que está describiendo la característica con el fin de facilitar su interpretación por parte del usuario. Por ejemplo, la característica ‘color’ se formaliza mediante el elemento *property* con la designación ‘tiene_color’. Así, hemos formalizado el rol descriptivo de ‘color’ en forma de *Object property*, hemos declarado que el dominio es ‘baldosa cerámica’ y que los valores de la propiedad son miembros de la clase ‘color’.

La formalización de las características en forma de *Object property* o de *Datatype property* se basa en el tipo de valor de la característica. En la identificación de los valores de la característica de la baldosa, hemos considerado las posibilidades de formalización en Protégé de esos valores y hemos distinguido previamente entre las palabras o términos que queremos incluir en forma de instancia o clase, los valores numéricos (decimales o enteros) y los valores booleanos. Las características con valores que son palabras o términos (o abreviaturas, siglas o símbolos) se han formalizado a través de *Object property* y las demás a través de *Datatype property*. La importancia de esta distinción radica en los conceptos que van a representarse en forma de clase o instancia en la ontología con respecto de aquellos que no es relevante formalizar en forma de clase o instancia. Por ejemplo, los posibles valores de la característica acabado son ‘esmaltado’ y ‘no esmaltado’ y en caso de querer aumentar el nivel de detalle, entre

²⁸ Para facilitar la comprensión del lector, iremos recordando qué son los elementos OWL que mencionamos a la vez que explicamos cómo los entendemos en nuestro trabajo.

los posibles valores de ‘no esmaltado’ encontramos ‘pulido’ y ‘sin pulir’. Estos conceptos resultan relevantes para describir la característica ya que constituyen su extensión y esta extensión tiene relevancia desde el punto de vista terminológico.

En el caso de los valores de las características representadas en forma de *Datatype* se trata de valores numéricos o booleanos que no interesa recoger en la clase de la característica en toda su extensión. En este caso, interesa determinar el tipo de valor que puede tomar la característica para así conocer más acerca de la extensión de ésta.

Independientemente de que el rol descriptor de las características de la baldosa cerámica se haya formalizado a través de *object property* o *datatype property*, todas estarán enlazadas a la clase que las representa a través de una anotación utilizando el subelemento ‘isDefinedBy’ que enlazara la propiedad con su clase de característica. Por ejemplo, en la figura 90 se muestra un fragmento de la ventana de descripción de la *object property* que representa el rol descriptor de la característica ‘acabado’. En esta figura, en la parte superior derecha que corresponde a la subventana de anotaciones, podemos observar el subelemento *isDefinedBy* indicando que esa propiedad se define en la clase ‘acabado’. Este tipo de anotación es distinto de otras anotaciones en lenguaje natural que iremos viendo en la ontología, aquí se ha formalizado la relación con una entidad de tipo clase y no se trata de una mera denominación en lenguaje natural. Lo podemos distinguir por el símbolo del rombo que precede a la clase e indica que se trata de un elemento formal presente en la ontología.

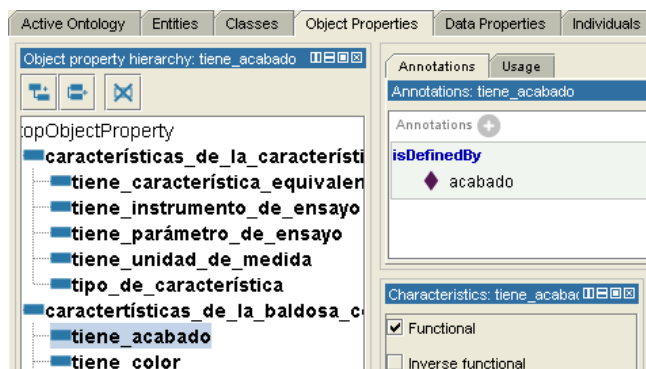


Figura 90. Vínculo de una propiedad que representa a la característica ‘acabado’ con la clase ‘acabado’

6.3.1 Características con valores de tipo unidad léxica

La formalización de los valores de tipo unidad léxica de las características se han realizado en forma de instancia de la clase de característica. Al declarar un valor como instancia de la clase estamos declarando que es un ejemplar de la clase, justo el aspecto que nos interesa representar de los valores de una característica. La instancia queda vinculada a la clase al haber sido asignada a la clase y la clase contendrá como miembros todas las instancias que se le hayan asignado. Así, por ejemplo, los valores identificados para la clase ‘color’ se han instanciado a la clase (figura 91).

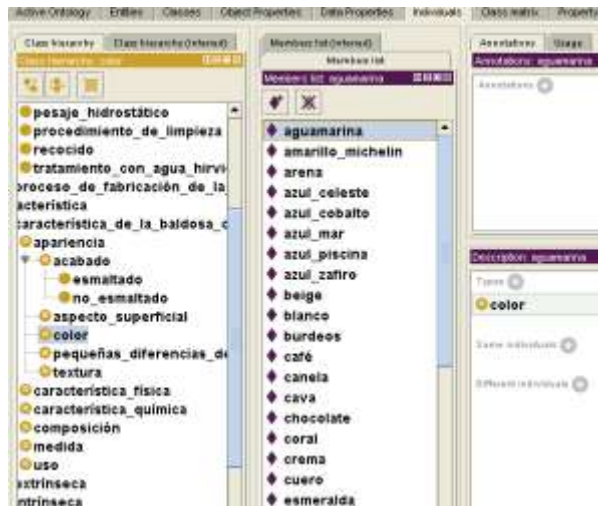


Figura 91. Instancias de la clase 'color' que representan los valores de 'color'

En la figura 84, observamos la pestaña *Individuals* donde aparecen, en el centro, algunos de los valores en forma de instancia que pertenecen a la clase color. A la izquierda se visualiza la clase a la que pertenecen los valores y a la derecha, el detalle de la instancia 'aguamarina' como tipo de 'color'.

Es importante remarcar que cuando restringimos el rango de una propiedad, estamos formalizado el tipo de valor que puede tener la característica. Así, por ejemplo, si en la descripción de una clase de baldosa hacemos uso de una propiedad y le asignamos un valor que no entra dentro del rango declarado en la propiedad, el razonador indicará que existe un error de inconsistencia. Por ejemplo, si describimos 'baldosa esmaltada' y declaramos 'tiene color:esmaltado', al no ser 'esmaltado' un valor posible para la propiedad color, el razonador indicará que existe una inconsistencia. Esto ocurre tanto en propiedades de tipo *object property* como en propiedades de tipo *datatype property*.

Todos los valores que eran palabras o términos se han formalizado de esta manera en la ontología. En el caso de que los valores tuvieran subvalores, los valores de primer nivel se han formalizado como subclase de la clase de característica y los valores de segundo nivel como instancias de las subclases. Por ejemplo, la característica 'acabado' tiene como posibles valores 'esmaltado' y 'no esmaltado'. Además, dentro de la posibilidad de que la baldosa tenga un acabado esmaltado, los valores posibles son 'alto brillo', 'brillo', 'mate', 'metalizado', 'nácar', 'rejunte' y 'satinado' y en caso de 'no_esmaltado' son 'pulido' y 'sin pulir'. La representación de los valores en la ontología pasa por declarar 'esmaltado' y 'no esmaltado' como subclases de 'acabado' (figura 92) y a continuación añadir las instancias correspondientes a la clase 'esmaltado' y a la clase 'no esmaltado' (figura 93 y 94).



Figura 92. Valores subclase de 'acabado'



Figura 93. Subvalores del valor 'esmaltado'

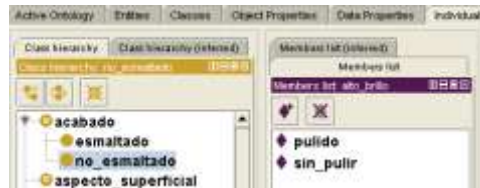


Figura 94. Subvalores del valor 'no esmaltado'

De esta forma se han formalizado todos los niveles de valor y se puede describir una baldosa utilizando un nivel de valor u otro según convenga.

Como hemos mencionado anteriormente, la formalización del rol descriptor de las características con valores de tipo unidad léxica se ha realizado a través de *Object property*. Cada *Object property* representa el rol descriptor de una característica y tendrá como dominio la clase de baldosa cerámica a la que describe o la clase de característica a la que describe, en el caso de características de las características, y como rango la clase de característica que incluye los valores posibles de ésta. Por ejemplo, la característica 'acabado' se ha formalizado a través de la *object property* 'tiene acabado', su dominio es la clase a la que describe, 'baldosa cerámica' y su rango la clase 'acabado' que agrupa los posibles valores de la característica (figura 95). En el caso de la característica 'tiene parámetro de ensayo', el dominio es 'características de la baldosa' y el rango 'parámetro de ensayo' (figura 95).

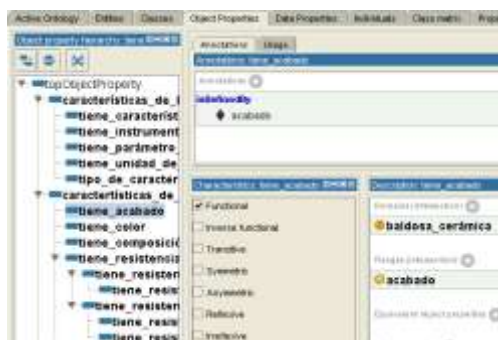


Figura 95. Formalización de la característica 'acabado' como *object property*

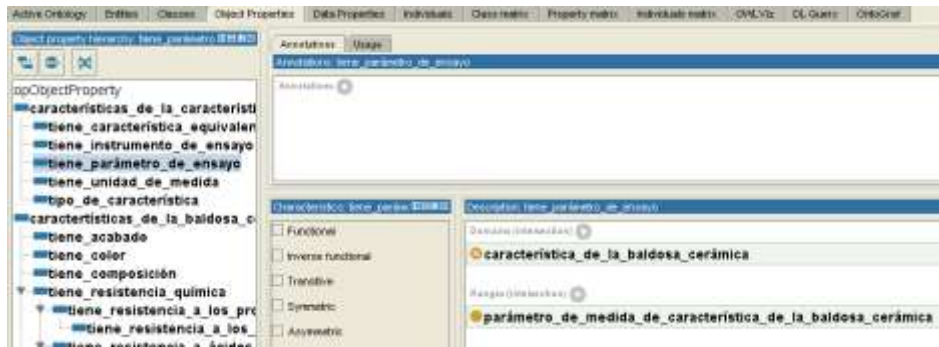


Figura 96. Formalización de la característica ‘parámetro de ensayo’ como *object property*

Para formalizar que una característica es funcional, es decir, sólo puede tomar un valor en la descripción con una misma clase, hemos indicado que la propiedad es funcional desde la ventana de descripción de la propiedad de Protégé (que equivale a declarar la *objectPropertyCharacteristic* 'Funcional'). En la figura 96, se puede observar la casilla *functional* marcada en la propiedad que representa a ‘acabado’.

6.3.2 Características con valores booleanos

La representación formal de una característica cuyo valor es booleano implica principalmente expresar que se trata de una característica cuyos valores son únicamente dos y que son opuestos (positivo y negativo). La representación formal de su potencial como característica booleana pasa por aprovechar las ventajas de procesamiento de las características representadas en forma de *datatype property* ya que, de esta forma, posteriormente podemos hacer consultas y los valores de este tipo de característica se procesa (calcula) utilizando un razonador capaz de interpretar que al ser un valor booleano, sus valores son opuestos.

La *datatype property* que representa el rol descriptor de una característica tendrá como dominio de la propiedad, la clase de concepto al que describe. En la ontología, las características formalizadas a través *datatype* describen a la clase ‘baldosa cerámica’, ‘baldosa esmaltada’ o ‘baldosa no esmaltada’. El rango de la propiedad se indica en la aplicación como valor de tipo *boolean*. En la figura 97 vemos un ejemplo de característica formalizada como *datatype property* con valor booleano.

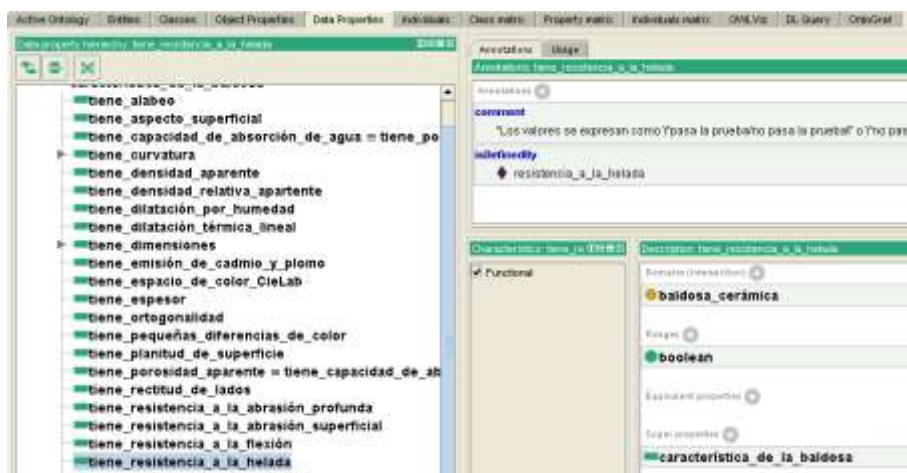


Figura 97. Formalización de ‘resistencia a la helada’ como *datatype property*

En la subventana *Description* de la propiedad ‘tiene resistencia a la helada’ visible en la imagen en la parte inferior derecha, se observa la declaración del dominio y el rango de la propiedad.

En la formalización de este tipo de características, hemos recogido también las expresiones en lenguaje natural que expresan el valor positivo o negativo de la característica. Estas expresiones se han formalizado mediante anotaciones utilizando el subelemento de anotación *comment*. En la figura 97, en la parte superior derecha vemos una anotación en forma de comment que indica “Los valores se expresan como "pasa la prueba/no pasa la prueba" o "no pasa el ensayo/pasa el ensayo”.

6.3.3 Características con valores numéricos sin rango limitado

En la fase de recogida de datos, hemos encontrado numerosas características que tienen como valores números (decimales o enteros). Dentro de este grupo de características, hay algunas cuyos valores posibles son limitados, por ejemplo, porque el número indica el tipo de grado de resistencia según un parámetro de ensayo y el número de grados posibles es limitado. Por otra parte, encontramos características cuyos valores son números pero no hay una restricción, *a priori*, sobre el valor máximo y mínimo que integra la clase de característica.

La formalización del rol descriptor de las características que no tienen un rango limitado se ha realizado en forma de *datatype property* y asignado como valor de la propiedad el elemento *double* (que representa valores que son números decimales) o *integer* (para valores que son números enteros). El dominio de estas propiedades, al representar características de la baldosa, será ‘baldosa cerámica’, ‘baldosa esmaltada’ o ‘baldosa no esmaltada’ según la característica en concreto.



Figura 98. Descripción de dominio y rango de ‘tiene alabeo’

En la figura 98 se visualiza la descripción de dominio y rango de la característica ‘alabeo’, el dominio es la clase ‘baldosa cerámica’ y el rango indica que los valores de esta característica serán números decimales. Como hemos mencionado anteriormente, al declarar que el valor es de tipo número decimal, hemos formalizado una restricción que deberá cumplirse a lo largo de toda la ontología. Por lo que, en el caso de ‘alabeo’, no se asignar otro tipo de valor diferente a un número decimal, por ejemplo, si el número a expresar fuera entero debería expresarse en su forma decimal (1.0 y no 1)

6.3.4 Características con valores numéricos con rango limitado

En este caso, la formalización de una característica de este tipo debe reflejar que los valores posibles de la característica estarán comprendidos en un rango limitado y que un

concepto descrito mediante una característica de este tipo debe tener un valor comprendido en el rango.

En esencia, este tipo de características coinciden con lo que hemos expresado anteriormente para características con valores numéricos ilimitados en la naturaleza de sus valores: se trata de características que expresan un valor cuantitativo a través de un número que indica el valor en referencia a una unidad de medida; pero además, este grupo se distingue porque las características tendrán valores que hacen referencia a una escala cerrada, acordada y no siempre vinculada a una unidad de medida (por ejemplo aspecto superficial).

La principal diferencia respecto a las características con valores numéricos no limitados radica en que el número o abanico de valores cuantitativos debe restringirse *a priori* para expresar adecuadamente la comprensión de la característica. La formalización de un rango con una restricción de valores numéricos se realiza utilizando la sintaxis de OWL 2 a través de la siguiente expresión: $datatype[>= número^{datatype}, <= número^{datatype}]$. Esta expresión puede declararse también de forma simplificada a través de la siguiente expresión $datatype[>= número, <= número]$ siempre que los números que utilicemos correspondan al *datatype* que hemos declarado (si es *double*, decimales, si es *integer*, enteros).

Por ejemplo, si en el rango de la *datatype property* que representa a ‘resistencia a la abrasión superficial’ declaramos: $integer[>= 0, <= 5]$, estamos formalizando que los números que pueden formar parte del valor de la propiedad son enteros y que concretamente, se debe tratar de un número entero mayor o igual a cero y menor o igual a cinco. Hemos declarado, entonces, que la extensión de valores de la característica ‘resistencia a la abrasión superficial’ es un número que puede ser cero, uno, dos, tres, cuatro o cinco.



Figura 99. Descripción de dominio y rango de ‘tiene resistencia a la abrasión superficial’

En la figura 99, se visualiza la descripción de dominio y rango de ‘tiene resistencia a la abrasión superficial’. El dominio, la representación de clase de concepto al que describe la característica, se ha declarado de la misma forma que con otras propiedades que hemos visto anteriormente. En este caso, se trata de ‘baldosa esmaltada’. De esta forma, se ha declarado que la característica ‘resistencia a la abrasión superficial’ no es una característica que describa a todos los tipos de baldosa sino únicamente a aquellas que se puedan agrupar bajo la clase ‘baldosa esmaltada’.

6.3.5 Característica con una expresión matemática compleja como valor

La dilatación térmica lineal de una baldosa se expresa a través de una fórmula matemática que se refiere al coeficiente de dilatación lineal de la baldosa tal y como se observa en el extracto de la norma UNE-EN 10545-8 (ISO, 1997d):

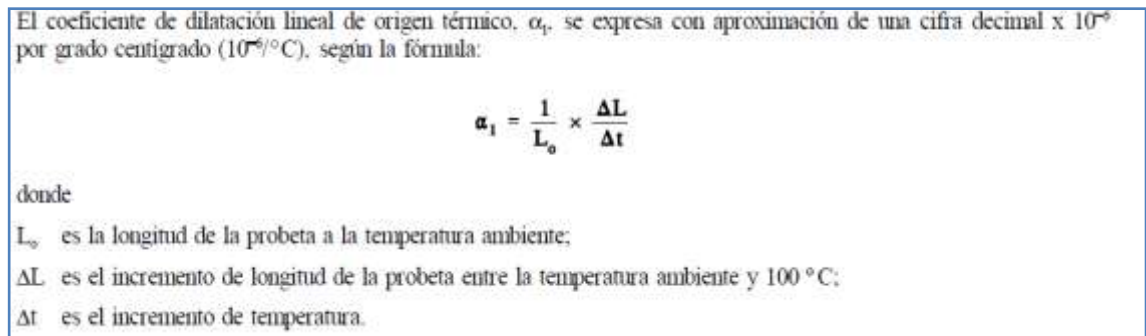


Figura 100. Extracto de la norma UNE-EN 10545-8

La representación forma de esta fórmula en OWL 2 supone, entre otros, un problema principal: no existe un tipo de valor definido para una *datatype property* que permita introducir el valor expresado para la característica.

La solución tomada ha sido utilizar el tipo de valor *string* que representa un valor de tipo expresión en lenguaje natural. La fórmula no será procesada como dato calculable, es decir, los valores matemáticos de la fórmula no se consideraran y pasaran a ser una suma de caracteres sin carga matemática. De este modo, en caso de querer representar la descripción de una baldosa con dilatación térmica de $7 \cdot 10^{-6}/^{\circ}\text{C}$ se visualizaría de la siguiente forma:



Figura 101. Ejemplo ficticio de uso de la propiedad de dilatación térmica lineal

La realidad es que en las clasificaciones de baldosas utilizadas, la dilatación térmica lineal no toma ningún valor en concreto por lo que no hemos podido aplicar esta técnica de representación en la ontología pero sí hemos formalizado la característica de modo que se pueda realizar en caso de necesitarlo.

6.3.6 Observaciones

A lo largo de la etapa de análisis para la formalización de características, hemos encontrado dos casos en los que la característica puede formalizarse de acuerdo con más de un tipo de valor, ‘aspecto superficial’ y ‘dilatación por humedad’.

En el caso de ‘aspecto superficial’, su rol descriptor se formaliza a través de una *datatype property* que podría tener valores de tipo *boolean* o *integer*. Los valores de tipo *boolean* hacen referencia a si una determinada baldosa cumple con el requisito de ‘aspecto superficial’ que es tener un ‘95%’ de baldosas libres de efectos. En el caso de los valores de tipo *integer* hacen referencia al porcentaje de baldosas libres de efectos.

La formalización del rango de esta característica sólo tomará en cuenta un tipo de valor. Con esto se pretende evitar que la descripción de un tipo de baldosa a través de esta propiedad

se realice de una forma y en otra de otra, lo que provocaría que no se pudiera inferir información sobre esas descripciones al considerarse elementos diferentes. Por ejemplo, si se describe una baldosa indicando que el ‘aspecto superficial’ de la baldosa es de ‘95’ y en otra se indica que la baldosa tiene ‘aspecto superficial’ (valor: true), el razonador no podría inferir que en realidad estas baldosas comparten la característica y podría ser equivalentes, o se necesitaría agregar una regla que indicara este fenómeno, aspecto que no vamos a considerar en este trabajo.

En el caso de ‘dilatación por humedad’, cuyos valores son de tipo *double* (números decimales), éstos pueden referirse a milímetros por metro o a porcentaje de expansión. La formalización del rango de esta propiedad, por tanto, debe dejar muy claro a qué tipo de valor se refiere; de otro modo, en una baldosa podríamos indicar un valor decimal referente a milímetros por metro y en otra a porcentaje de expansión y parecería que son el mismo valor. Nuevamente, la formalización de este tipo de característica se ha realizado restringiendo a un único tipo el valor posible del rango de la propiedad.

No obstante, esta información sobre otros valores posibles se ha reflejado en la ontología en forma de anotación. A continuación, describimos la información que se ha formalizado a través de anotaciones pero a modo de muestra, en la figura 102, se observa en la sub-ventana ‘Annotations’ cómo hemos reflejado la existencia de un valor alternativo en forma de anotación de tipo *comment* y cómo indicamos el valor de referencia a través del elemento *SeeAlso*.

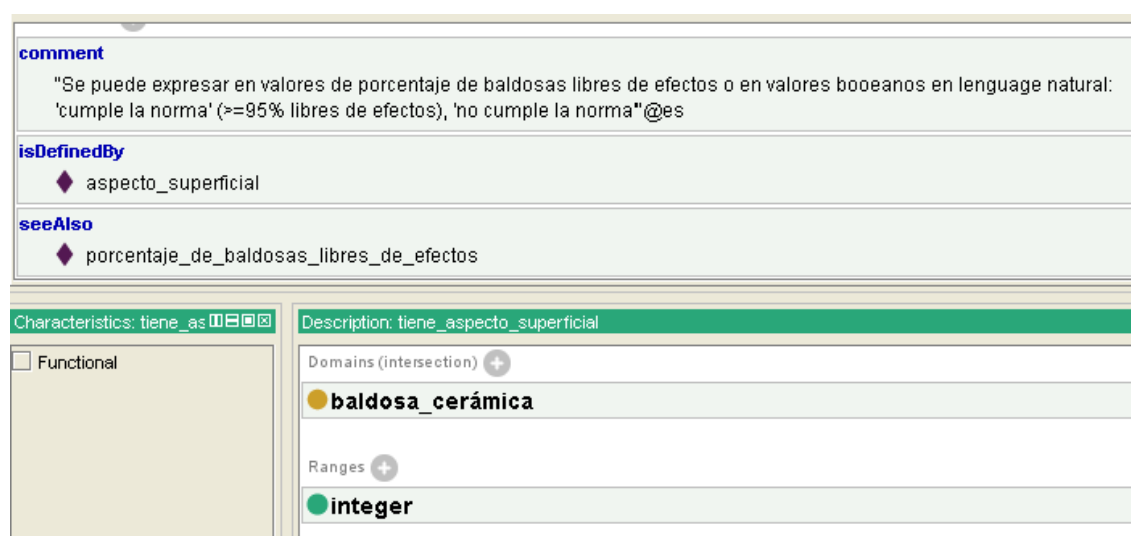


Figura 102. Anotaciones y descripción del dominio y rango de la propiedad de ‘aspecto superficial’

6.4 Formalización de características equivalentes

La formalización de las características que describen las características de la baldosa se ha realizado, como hemos indicado anteriormente, a través de *object property*. Sin embargo, en el caso de la descripción de características equivalentes, la formalización ha considerado la representación de la equivalencia computacional y no sólo descriptiva entre las características implicadas tal y como explicamos a continuación.

Las características equivalentes en la ontología son ‘capacidad de absorción de agua’ y ‘porosidad aparente’. La formalización de la equivalencia entre estas dos características se ha

realizado a través de una *object property* denominada ‘tiene característica equivalente’ que tiene como dominio y rango ‘característica de la baldosa cerámica’ y que es simétrica.

En la clase ‘capacidad de absorción de agua’ se han introducido dos expresiones OWL 2 en el correspondiente apartado *Superclasses* que refleja que la característica tiene una característica equivalente y que la característica equivalente es ‘porosidad aparente’. Del mismo modo, en la clase ‘porosidad aparente’ se han introducido dos expresiones OWL 2 en el correspondiente apartado *Superclasses* que refleja que la característica tiene una característica equivalente y que la característica equivalente es ‘capacidad de absorción de agua’ (figura 103).



Figura 103. Descripción en las clases ‘capacidad de absorción de agua’ y ‘porosidad aparente’ de la equivalencia entre ellas

La relación de equivalencia, tal y como la entendemos, se da en la característica cuando se observa como rol descriptor y no entre las características como conceptos, ya que tienen una intensión y extensión distintas. Por eso, esta restricción conlleva que formalicemos también la equivalencia entre las propiedades que representan a cada característica y no en las clases ‘porosidad aparente’ y ‘capacidad de absorción de agua’. La equivalencia entre las propiedades ‘tiene capacidad de absorción de agua’ y ‘tiene porosidad aparente’ se formaliza desde la pestaña *Data properties* indicando en el apartado *Equivalent properties* de alguna de las dos propiedades que la otra es equivalente.

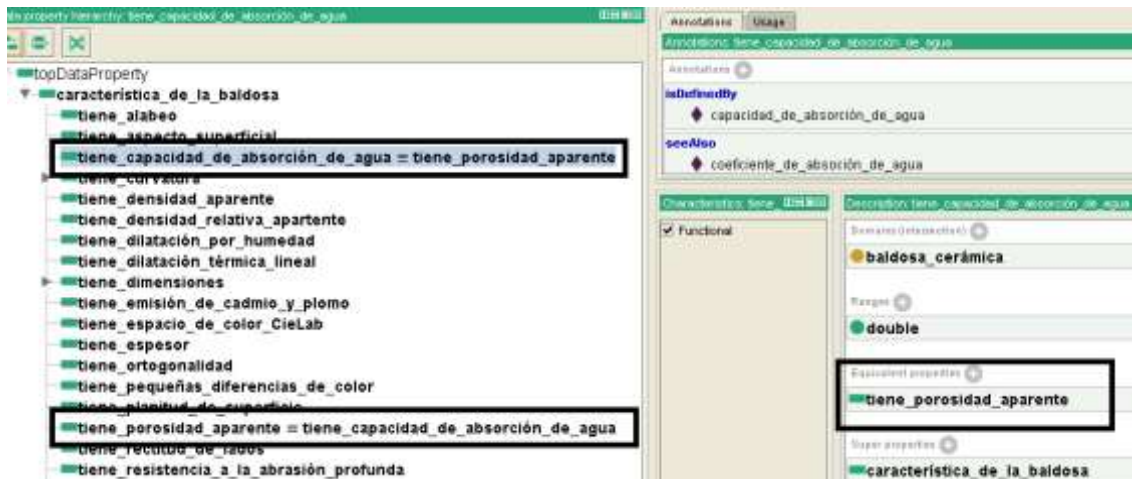


Figura 104. Inferencia en la propiedad ‘tiene capacidad de absorción de agua’ de la propiedad equivalente ‘tiene porosidad aparente’.

La figura 104 muestra la formalización de ‘tiene porosidad aparente’ como equivalente de ‘tiene capacidad de absorción de agua’, al pasar el razonador se ha inferido esta misma información sobre la equivalencia en la descripción de ‘tiene porosidad aparente’.

6.5 Formalización de otra información en lenguaje natural

Hasta ahora hemos visto que las características se han formalizado a través de clases y se han estructurado en la ontología junto con el resto de conceptos implicados en su descripción. También hemos visto que el rol descriptor de las características se ha formalizado a través de propiedades y que estas propiedades están vinculadas a las clases a través de una anotación que apunta a la clase con el elemento “*IsDefinedBy*”.

En la ontología también hemos expresado en forma de anotación otros aspectos a considerar sobre las propiedades con respecto a la expresión de valores, el parámetro de ensayo utilizado o la unidad de medida. Esta información aparece formalizada en la clase de característica pero hemos considerado de utilidad representarla también en la propiedad que presenta a cada característica. Estas anotaciones, que actúan como información complementaria para describir la propiedad, se han formalizado utilizando “*SeeAlso*”. Por ejemplo, en la figura 105, se visualizan las anotaciones de la propiedad que representa a ‘resistencia a la abrasión superficial’, además de incluir el enlace a la clase de característica donde se describe esta característica, se muestra el elemento *SeeAlso* apuntando a la clase ‘defecto visible a revoluciones’ de la ontología, el parámetro utilizado para medir esta característica.

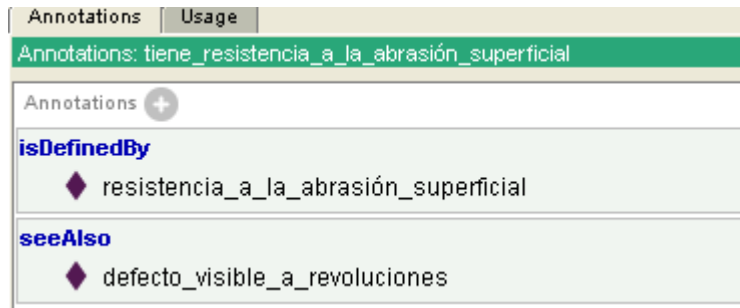


Figura 105. Anotaciones de la propiedad de ‘resistencia a la abrasión superficial’

Este subelemento se ha utilizado también para distinguir la referencia a la fuente donde hemos encontrado la característica de otras fuentes complementarias, la referencia a la fuente se ha formalizado utilizando el subelemento *seeAlso* en la clase de característica.

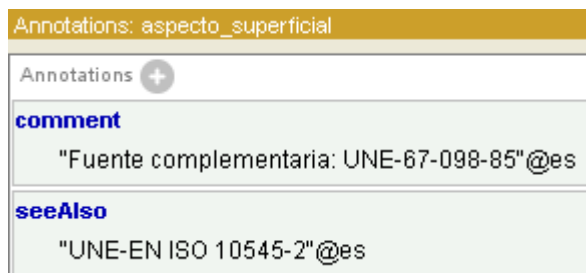


Figura 106. Anotación sobre la fuente de la clase ‘aspecto superficial’ y sobre otra fuente relevante

Las referencias complementarias, los contextos y los comentarios sobre las clases de las características se han formalizado utilizando el subelemento *comment*. La representación de los contextos a través de anotaciones con el subelemento *comment* se ha realizado siguiendo el patrón “contexto_número de contexto: contexto [fuente de contexto]”.

Annotations: aspecto_superficial

comment

"Contexto1: Los criterios para valoración del aspecto superficial de las baldosas esmaltadas, engobadas o no esmaltadas, son los siguientes:

- a) Grietas
- b) Cuarteado
- c) Falta de esmalte
- d) Ondulado
- e) Cráteres
- f) Agujeros
- g) Desvitrificación del esmalte
- h) Puntos y manchas.
- i) Defectos bajo el esmalte.
- j) Defectos de decoración.
- k) Destonificación.
- l) Astillado de cantos
- m) Despuntado.

Para poder juzgar si un efecto decorativo intencionado es aceptable o hay un defecto, ver la correspondiente cláusula de Norma de Producto. Los defectos a), l) y m), no pueden ser efectos intencionados. [Ce005-7e.txt]

Contexto 2: Esta parte de la Norma ISO 10545 establece métodos para la determinación de las características dimensionales (longitud, anchura, grosor, rectitud de lados, ortogonalidad, y planitud de superficie), y de la calidad superficial de las baldosas cerámicas.

Las baldosas con área inferior a 4 cm² no son objeto de medida de longitud, anchura, rectitud de lados, ortogonalidad, y

Figura 107. Ejemplo de contextos de 'aspecto superficial'

La formalización de observaciones sobre otros conceptos también se ha formalizado utilizando los mismos elementos, a continuación, se muestran dos ejemplos de observaciones, en la instancia que representa el valor de una característica (figura 108) y en un parámetro de medida de la característica (figura 109).

Annotations: pasta_roja

Annotations +

comment

"pasta compuesta de arcillas que, durante el proceso de cocción, adquieren coloración que va desde el amarillo paja hasta el rojo o marrón intensos, en función del contenido de óxido de hierro y, en menor medida, óxido de manganeso. [PROALSO, 2011]"@es

Figura 108. Anotación en el valor 'pasta roja' de la característica 'composición'

Annotations: dureza_de_Mohs

Annotations +

comment

"1 Talco Se raya fácilmente con la uña
2 Yeso Se raya con la uña
3 Calcita Se raya con moneda de cobre
4 Fluorita Se raya fácilmente con lámina de acero
5 Apatito Se raya dif. con lámina de acero
6 Feldespato Se raya con lima metálica
7 Cuarzo Dif. con lima metálica
8 Topacio Raya el vidrio de ventana
9 Corindón Raya el vidrio de ventana
10 Diamante Raya el vidrio de ventana
[CE020-2e.txt]"@es

Figura 109. Anotación sobre el parámetro 'dureza de Mohs'

En el caso de anotaciones sobre las características en las propiedades (object property o datatype property) de la ontología, éstas se formalizan utilizando el subelemento *comment*. Por

ejemplo, en la figura 110, vemos una explicación de los posibles valores de ‘tiene aspecto superficial’ en forma de anotación con el subelemento *comment*.

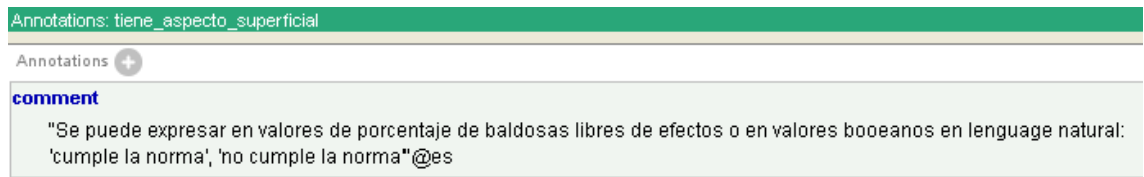


Figura 110. Observación sobre los valores posibles de ‘tiene aspecto superficial’ en forma de anotación con el subelemento *comment*

6.6 Formalización de las relaciones en la ontología

La formalización de las relaciones se ha realizado siguiendo los mismos principios que inspira la formalización del rol descriptor de las características: a través de propiedades. Las relaciones representadas en la ontología vinculan una clase de concepto con otra clase de concepto, de manera que el dominio de la propiedad es una clase de concepto, la etiqueta de la propiedad refleja la relación existente y el rango de la propiedad corresponde a la clase de concepto con la que el concepto del dominio mantiene la relación.

En este caso la formalización de relaciones siempre pasa por la declaración de información en forma de *object property*.

6.6.1 Relación baldosa cerámica – proceso de fabricación

La relación baldosa cerámica – proceso de fabricación refleja el vínculo existente entre un tipo de baldosa cerámica y el proceso utilizado para la fabricación de la baldosa. Esta relación se ha formalizado a través de una *object property* cuyo dominio es la clase ‘baldosa cerámica’ y cuyo rango es la clase ‘proceso de fabricación’.



Figura 111. Formalización de la relación baldosa cerámica – proceso de fabricación

En la figura 111, se muestra la ventana de formalización de esta relación. En la parte superior derecha, visualizamos una anotación sobre la clase donde se describen los procesos relacionados con esta relación. En la parte inferior derecha, en la ventana de *Description*, visualizamos el dominio y rango declarados para esta relación. Igualmente, a la izquierda de esta ventana, se visualiza cómo hemos declarado que se trata de una relación de tipo funcional.

6.6.2 Relación característica – proceso de ensayo

La relación característica – proceso de ensayo refleja el vínculo existente entre una característica y el proceso de ensayo utilizado para determinar los valores de la característica. Esta relación se ha formalizado a través de la *datatype property* ‘relación característica – proceso’ y tiene como dominio la clase ‘característica de la baldosa cerámica’. El rango de la propiedad es ‘proceso de ensayo de característica’ donde se agrupan los distintos procesos relacionados con el ensayo de los valores de una característica. En la figura 112, se visualiza la descripción de esta propiedad.



Figura 112. Formalización de la relación característica - proceso

6.6.3 Relación de meronimia

La relación de meronimia refleja el vínculo entre una característica y otra que es parte constitucional de la anterior. En este caso la formalización de esta relación se ha realizado a través de una *datatype property* y una expresión en la clase de característica.

La propiedad que representa esta relación tiene la misma clase como dominio y rango, la clase ‘característica de la baldosa’. En la figura 113, se muestra la descripción de dominio y rango de la propiedad.



Figura 113. Dominio y rango de la propiedad para la relación de meronimia

No obstante, esta relación también se expresa en la clase de la característica que compone el todo. La formalización de esta relación a través de la clase implica dos declaraciones y la descripción de las clases implicadas. En primer lugar, en la clase que representa al todo se declara que ésta es equivalente al conjunto de las partes. Por ejemplo, en la figura 114, se muestra la declaración de clases de ‘planitud de superficie’ y su equivalencia a la clase ‘alabeo’ o a la clase ‘curvatura’.



Figura 114. Declaración de clases parte de la clase ‘planitud de superficie’

Además, en la clase ‘alabeo’ y en la clase ‘curvatura’ se ha declarado que tienen como superclase la clase ‘planitud de superficie’ (figura 115).



Figura 115. Declaración de la superclase de ‘curvatura’ y ‘alabeo’

Esta es la forma de expresar OWL 2 que existe una relación de meronimia entre ‘planitud de superficie’ y ‘alabeo’ y ‘curvatura’. La clase ‘alabeo’ y la clase ‘curvatura’ son un tipo de clase de ‘planitud de superficie’ pero un tipo complementario donde cada clase cubre un espacio de la clase ‘planitud de superficie’ (ya que hemos indicado que ‘planitud de superficie’ es equivalente a una clase o la otra).

6.7 Formalización del tipo de característica intrínseca o extrínseca

El hecho de identificar si una característica es intrínseca o extrínseca en Terminología ayuda en la tarea de categorizar conceptos y en la de identificar relaciones entre el concepto descrito y otros conceptos. Por eso, resulta interesante no sólo añadir esta información en cada una de las características de forma independiente sino también poder clasificarlas de forma automática en forma de clases (intrínsecas o extrínsecas) de acuerdo con este criterio. Dado que en una ontología podemos realizar tantas jerarquías como queramos sin que ello impida que una característica se pueda situar bajo distintas clases (siempre que no resulte incoherente), este aspecto no resulta un impedimento.

Como la descripción de si una característica es intrínseca o extrínseca forma parte de la información relativa a la naturaleza de la característica independientemente del concepto al que pueda describir, consideramos que corresponde a la característica como clase acoger esta información. Para ello, debemos añadir esta información en la descripción sobre las características en su descripción como clase en la ontología que es donde la característica se ve formalizada como concepto.

La formalización se ha realizado conjugando la descripción de las clases de característica de la baldosa con la descripción de las clases que representan ‘característica intrínseca’ y ‘característica extrínseca’ subclases de la clase ‘característica’.

En primer lugar, se ha formalizado a través de una *object property* que una característica de la baldosa puede ser de tipo extrínseco o intrínseco. Para esto, hemos indicado que el dominio de la propiedad es la clase ‘característica de la baldosa cerámica’. En el rango de la propiedad hemos declarado que sólo puede tomar dos valores ‘característica intrínseca’ o ‘característica extrínseca’ y hemos reflejado que sólo puede ser uno u otro, y no ambos, a través de la primitiva *or* (figura 116).



Figura 116. Dominio y rango de la propiedad que representa el tipo de característica

El resto de la formalización se realiza en las clases. En el caso de las características de la baldosa, se indica que la clase tiene la propiedad ‘tipo de característica’ y a través de la expresión *some*, en el apartado de *Superclasses* de la característica, se declara que el valor es ‘característica extrínseca’ o ‘característica intrínseca’ según corresponda.

En la figura 117, mostramos la descripción de la clase ‘apariencia’ como característica de tipo intrínseco. Entre otras descripciones, vemos como se ha declarado que apariencia pertenece al conjunto de clases que tienen como ‘tipo de característica’ el valor ‘característica intrínseca’.

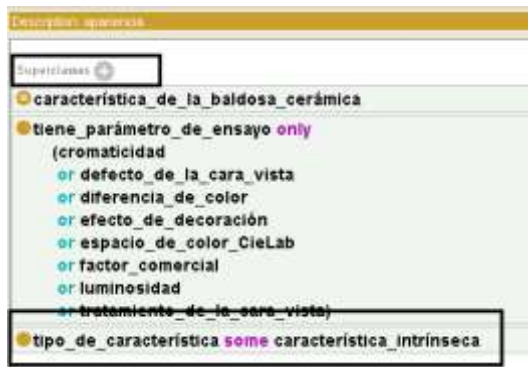


Figura 117. Descripción de la clase ‘apariencia’ como característica intrínseca

Es necesario remarcar que las subclases de una clase descrita como ‘característica extrínseca’ o como ‘característica intrínseca’ heredan esa descripción, por lo que no es necesario formalizar esta información de nuevo en las subclases.

Finalmente, la formalización termina declarando en la clase ‘característica intrínseca’ que las clases que tengan la propiedad ‘tipo de característica’ con valor ‘característica intrínseca’ pertenecen a esa clase desde *Equivalent classes* y a través del elemento *some* (figura 118). Del mismo modo, en la clase ‘característica extrínseca’ se ha declarado que las clases que tengan la propiedad ‘tipo de característica’ con valor ‘característica extrínseca’ pertenecen a esa clase (figura 119).



Figura 118. Descripción de la equivalencia en la clase ‘característica intrínseca’



Figura 119. Descripción de la equivalencia en la clase ‘característica extrínseca’

Así, hemos formalizado la clase ‘característica extrínseca’ es equivalente a las clases cuyo tipo de característica es ‘característica extrínseca’ y que la clase ‘característica intrínseca’ es equivalente a las clases cuyo tipo de característica es ‘característica intrínseca’. Dado que en las descripciones de las clases de característica de la baldosa hemos indicado que tienen como superclase una clase cuyo tipo de característica es ‘característica intrínseca’ o ‘característica extrínseca’, estas clases de características de la baldosa cerámica se considerarán subclases de ‘característica extrínseca’ o ‘característica intrínseca’ según se haya declarado.

Esta opción no era la única para formalizar esta información en la ontología, podría haber bastado con declarar que la superclase de una característica de la baldosa era ‘característica intrínseca’ o ‘característica extrínseca’ según correspondiera. No obstante, esta opción permite reflejar que se trata de una tipología (característica intrínseca o extrínseca) distinta a la tipología de clases de características de la cerámica general que se basa en el análisis del ámbito (apariciencia, característica física, característica química, composición, medida y uso) e incluye la etiqueta ‘tipo de característica’ para reflejar esa distinción.

6.8 Formalización de características esenciales y distintivas en la ontología

La formalización de una característica como esencial o distintiva en la ontología está orientada a describir las clases de característica de la baldosa cerámica y las clases de baldosa cerámica. A continuación describimos los enfoques y los elementos que hemos utilizado para formalizar estas clases de conceptos, distinguiendo entre la formalización de las clases de característica de la baldosa cerámica y la formalización de clases de baldosa cerámica.

Esta distinción se debe a que en cada formalización perseguimos un objetivo distinto, en el caso de las características el objetivo es por una parte probar si la clasificación inicial de características es consistente. Es decir, si a partir de la descripción realizada, el razonador infiere la misma clasificación que hemos representado manualmente, sin tener en cuenta en la comparación la distinción entre característica intrínseca y extrínseca. Por otra parte, persigue probar si la descripción formal de cada clase es suficientemente precisa como para que cada una de las características se distinga del resto y se agrupe bajo la clase o clases correcta/s.

En el caso de las baldosas cerámicas, la clasificación inicial no responde a una jerarquía consistente. Por ejemplo, el primer nivel de la jerarquía tiene como tipos de baldosa: ‘baldosa de gres rústico’, ‘baldosa no esmaltada’ o ‘grupo ISO de baldosa’. Un análisis de las diferencias y

semejanzas entre estas clases de baldosa sugiere que la característica distintiva entre estas hermanas de clase no es una misma sino una u otra dependiendo de entre quien se comparen las hermanas de clase e incluso, unas son subclases de otras. Por ejemplo, entre ‘baldosa no esmaltada’ y ‘baldosa de gres rústico’ no existe una igualdad, ya que ‘baldosa de gres rústico’ es un tipo de ‘baldosa no esmaltada’.

Esta clasificación persigue probar otro de los objetivos de la tesis, describir las clases de baldosa a través de las características de la ontología y probar si la *reclasificación* de baldosas cerámicas con un razonador es consistente y da como resultado una estructura coherente con los datos extraídos sobre cada clase de baldosa cerámica. En otras palabras, probar si las características que hemos formalizado resultan suficientes para describir las baldosas cerámicas y si la clasificación basada en estas características se corresponde con la realidad.

6.8.1 Descripción de las clases de característica de la baldosa cerámica

La formalización de las características y relaciones de una clase de característica de la baldosa cerámica se ha realizado considerando un enfoque *bottom-up* de descripción de las clases. Es decir, la descripción de las clases tiene en cuenta la jerarquía de características representada en la ontología y describe a las características clasificadas en el nivel más bajo de la jerarquía primero para pasar a la descripción del nivel superior a continuación hasta llegar al nivel de la clase ‘característica de la baldosa cerámica’. Esta descripción no se aplica a clases de la jerarquía que representan valores de las características. Por ejemplo, en la figura 120, se muestra la jerarquía de clases de la clase ‘apariencia’, la formalización de características esenciales y distintivas se realiza en las clases ‘textura’, ‘acabado’, ‘color’, ‘aspecto superficial’ y ‘pequeñas diferencias de color’ y no se consideran las clases ‘esmaltado’ o ‘no esmaltado’ al ser valores de la clase ‘acabado’. Recordemos que ‘esmaltado’ y ‘no esmaltado’ se han formalizado como clases para poder agrupar otros valores más concretos y permitir descripciones más generales o más concretas (§III.6.3.1).

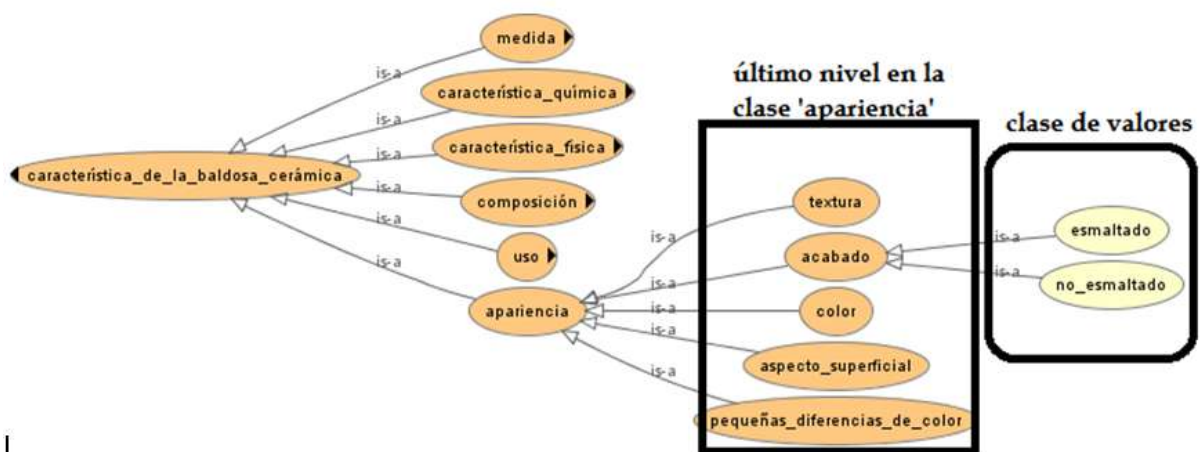


Figura 120. Jerarquía de clases de ‘apariencia’

La característica **distintiva** para cada una de las subclases de una clase será la misma, siguiendo el enfoque utilizado en sistemas como CAOS y se determinará de acuerdo con la idea de *diferencia específica* de la Ontotermología (§II.1.5), identificando cuál es la característica o

características que permite diferenciar la clase del resto de conceptos. En este sentido, la característica distintiva (o características distintivas) de una clase deberá ser la misma para todas las hermanas de clase, cambiando únicamente su valor. Esto probará que la clase es consistente y no contiene miembros que deberían pertenecer a otras subclases o clases. Si esto no ocurre, se entenderá que la clase no es consistente.

Por ejemplo, la característica que se ha formalizado como distintiva de ‘acabado’, ‘aspecto superficial’, ‘color’, ‘pequeñas diferencias de color’ y ‘textura’ entre ellas es ‘parámetro’, aunque además haya otras diferencias entre algunas de las clases indicadas. Por ejemplo, ‘acabado’, ‘aspecto superficial’, ‘textura’ y ‘pequeñas diferencias de color’ tienen instrumentos de ensayo distintos entre ellas, pero esta característica no es la que las distingue a todas las hermanas de clase’.

La formalización de la característica distintiva de una clase se realiza a través de la combinación de una expresión OWL 2 en el apartado de *Equivalent Classes* de la característica y otra en el apartado de *Super classes*. La expresión OWL 2 a través de *some* en *Equivalent Classes* declara que la clase equivale a tener la propiedad determinada con un determinado valor y la expresión a través de *only* en *Super Classes* indica que el valor de la propiedad sólo puede ser el indicado. Por ejemplo, en la figura 121, se muestra la descripción de la característica distintiva de la clase ‘acabado’ donde hemos formalizado la característica distintiva ‘tiene parámetro de ensayo’: ‘tratamiento de la cara vista’.



Figura 121. Fragmento de la descripción de ‘acabado’ con la característica distintiva ‘tiene parámetro de ensayo’

En el apartado de *Equivalent classes*, hemos declarado la expresión “tiene parámetro de ensayo *some* tratamiento de la cara vista” que representa que la clase ‘acabado’ equivale a las clases que tienen como parámetro de ensayo el tratamiento de la cara vista. Además, hemos indicado en *Superclasses* la expresión “tiene parámetro de ensayo *only* tratamiento de la cara vista” que indica que la clase ‘acabado’ sólo puede tener ‘tratamiento de la cara vista’ como valor de ‘tiene parámetro de ensayo’. De esta manera distinguimos la característica del resto de características que tendrán otros valores de la propiedad ‘tiene parámetro de ensayo’.

Las características **esenciales** y las relaciones de la clase que no sean compartidas por la superclase o que tomen un valor más restringido se han formalizado en el *Superclasses*. Es decir, si la característica es esencial para la clase pero comparte el mismo valor con todas sus hermanas, esta característica se formaliza en la clase superior. En cambio, si la característica es esencial y además toma un valor distinto de los valores atribuidos a alguna de sus hermanas, la formalización de esta descripción se realiza en la clase correspondiente, en el apartado de *Superclasses*.

La formalización se declara combinando una expresión OWL 2 que vincula la propiedad con un valor a través de *some* y una expresión que vincula la misma propiedad a través de *only*. De esta manera, indicamos que la característica de la baldosa tiene una característica o relación y que la característica o relación tiene un valor en concreto.

En la figura 122, se muestra un ejemplo de característica esencial formalizada en la clase ‘aspecto superficial’. En el apartado de *Superclasses* se ha declarado la expresión ‘tiene instrumento de ensayo *some* (fotómetro *and* luz fluorescente *and* regla)’ que indica que la clase tiene la característica instrumento de ensayo y que los instrumentos son ‘fotómetro’, ‘luz fluorescente’ y ‘regla’. Al mismo tiempo, la expresión ‘tiene instrumento de ensayo *only* (fotómetro *and* luz fluorescente *and* regla)’ delimita la posibilidad de valores a únicamente ‘fotómetro’, ‘luz fluorescente’ y ‘regla’.

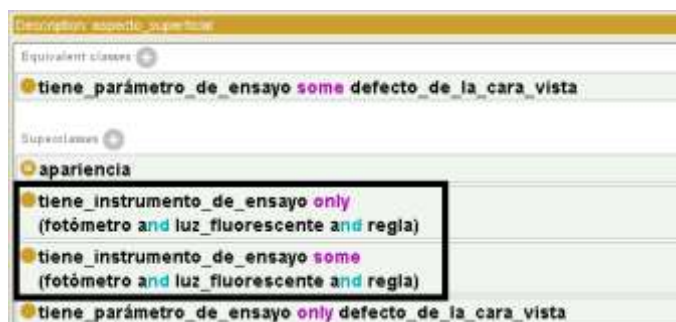


Figura 122. Extracto de la descripción de ‘aspecto superficial’ con una característica esencial delimitada

Las características de la clase de característica de la baldosa cuyos valores sean los mismos para todas las hermanas de clase, se han formalizado en la clase superior. Estas características se formalizarán como esenciales o distintivas dependiendo del resultado de la comparación de características con sus hermanas de clase. Por ejemplo, la clase *apariciencia* (superclase de ‘acabado’, ‘aspecto superficial’, ‘color’, ‘pequeñas diferencias de color’ y ‘textura’) se distingue de sus hermanas de clase (‘característica física’, ‘característica química’, ‘composición’, ‘medida’ y ‘uso’) nuevamente por su parámetro de ensayo. La descripción de la característica distintiva de clase ‘apariciencia’ incluye, en el apartado de *Equivalent classes*, la expresión:

tiene_parámetro_de_ensayo *some* (cromaticidad *or* defecto_de_la_cara_vista *or* diferencia_de_color *or* efecto_de_decoración *or* espacio_de_color_CieLab *or* factor_comercial *or* luminosidad *or* tratamiento_de_la_cara_vista)

Y se completa, en el apartado de *Superclasses* con la expresión:

tiene_parámetro_de_ensayo *some* (cromaticidad *or* defecto_de_la_cara_vista *or* diferencia_de_color *or* efecto_de_decoración *or* espacio_de_color_CieLab *or* factor_comercial *or* luminosidad *or* tratamiento_de_la_cara_vista)

El parámetro de ensayo de la clase ‘apariciencia’, que abarca todos los valores posibles de las subclases, distingue a la clase de la clase ‘apariciencia’ de sus hermanas de clase (figura 123).



Figura 123. Extracto de la descripción de la clase ‘apariencia’ con la característica distintiva ‘tiene parámetro de ensayo’

6.8.2 Descripción de las clases de baldosa cerámica

La descripción de las clases de baldosa cerámica también parte de un enfoque *bottom-up* de manera que se describen primero las clases que se encuentran en el último nivel de la jerarquía y cuando la descripción está completada se pasa al siguiente nivel. Coincide también con lo planteado para la descripción de características de la baldosa en formalizar las descripciones de características cuyos valores comparten todas las hermanas de una clase en la clase superior. Igualmente, los formalismos utilizados para describir la característica distintiva y las características esenciales o las relaciones son los mismos que los presentados en el apartado anterior. La identificación de las características distintivas y esenciales se basa en la comparación de las descripciones de las baldosas en las fichas de descripción de las baldosas y en el análisis general de la información recogida en cada ficha.

Una de las diferencias con respecto al método de formalización expresado en el apartado anterior radica en que la característica distintiva no siempre es la misma para todas las hermanas de una clase. Por ejemplo, la característica distintiva de la clase ‘baldosa esmaltada’ es el acabado de la baldosa, pero la clase ‘baldosín catalán’ (hermana de clase de ‘baldosa esmaltada’ en la jerarquía manual) se distingue de sus hermanas de clase por suma de las características acabado y la composición y la relación con el proceso de fabricación utilizado.



Figura 124. Descripción de la clase ‘baldosa esmaltada’ y su característica distintiva

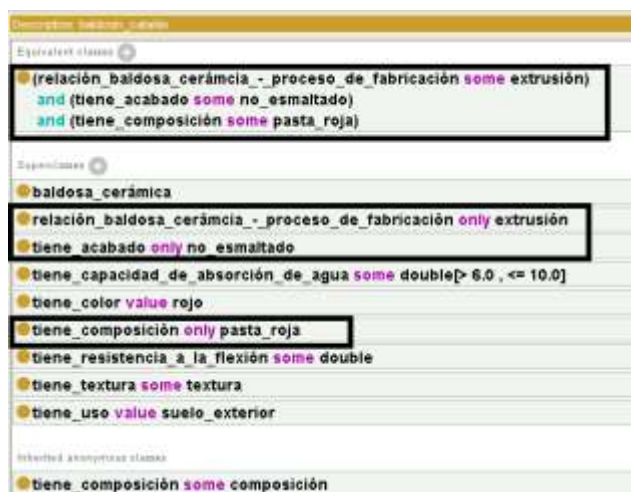


Figura 125. Descripción de la clase ‘baldosín catalán’ y sus características distintivas

La otra diferencia en la formalización de la descripción de las clases de baldosa cerámica se encuentra en la formalización de las características esenciales o la relación esencial de las clases. En este caso, la formalización de la característica de la baldosa se describe únicamente a través de la expresión OWL 2 con *some* en *Superclasses*, sin restringir que es necesario que tenga la propiedad. Así, formalizamos que la clase puede tener esa característica esencial con el valor declarado pero admite como miembros de la clase, baldosas que no tengan declarada esa característica. Por ejemplo, en la clase ‘grupo A’ las características esenciales son:

- tener un acabado esmaltado y resistencia a las manchas con un valor mínimo de 3 o no tener acabado esmaltado
- tener un acabado esmaltado y resistencia a los productos domésticos de limpieza con la clase GB de valor o tener un acabado esmaltado y resistencia a los productos domésticos de limpieza con la clase UB de valor
- tener un aspecto superficial mayor o igual a 95
- tener una ortogonalidad con un valor entre -1 y +1
- tener resistencia a la cuarteo

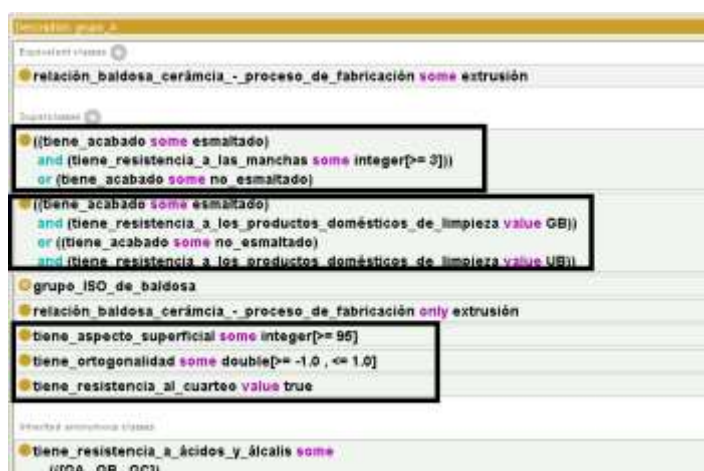


Figura 126. Descripción de la clase de baldosa cerámica ‘grupo A’ con las característica esenciales remarcadas

No obstante, al no haber restringido que estas propiedades deben aparecer en la descripción de sus subclases, la clase ‘grupo A’ admite como miembros de su clase, baldosas cerámicas que no incumplan estos criterios, aunque no tengan declaradas todas y cada una de las características descritas en la superclase. Es decir, si una baldosa tuviera alguna de estas características con un valor diferente no podría pertenecer a la clase, pero en cambio, si en la descripción de la baldosa no se ha indicado nada acerca de la propiedad, podría considerarse subclase de ‘grupo A’ siempre que cumpla con las características distintivas necesarias para pertenecer a ‘grupo A’. En este caso, cualquier baldosa cuyo proceso de fabricación fuera ‘extrusión’ podría pertenecer a esta clase si no contradice ninguna de las otras características esenciales declaradas.

Esta estrategia de formalización permite agrupar baldosas cerámicas que si bien, pudieran cumplir con los criterios de una superclase en cuanto a características distintivas, no tienen declarada en sus clases información relativa a otras características esenciales. En este caso, donde las descripciones que hemos identificado en los recursos textuales no siempre son exhaustivas y a menudo omiten características (lo que no quiere decir que no las tengan), este tipo de formalización permite identificar posibles superclases de baldosas que de otra forma no podrían inferirse.

6.9 Formalización de la sinonimia y equivalentes a otras lenguas

La sinonimia, entendida como otras denominaciones en español de algún elemento de la ontología, y las equivalencias, entendidas como otras denominaciones en inglés de algún elemento de la ontología, se han formalizado de manera que no afecta a la descripción de las clases y propiedades de la ontología. Es decir, que si bien es cierto que hemos utilizado el castellano como lengua de trabajo para la formalización de datos, la ontología tal y como está construida permite la adhesión de otras denominaciones en castellano o en otras lenguas sin tener que duplicar las clases descritas o repetir las descripciones formalizadas para cada lengua.

La formalización de sinónimos y equivalentes en esta ontología se ha centrado únicamente en las clases de característica de la baldosa y en las propiedades que las representan. Este trabajo no está orientado a la extracción de equivalentes en otras lenguas de conceptos de la cerámica y lo que hemos hecho ha sido recoger una pequeña muestra de equivalentes a partir de las normas ISO. Esta muestra tiene como finalidad probar la posibilidad de formalización de sinónimos y equivalentes en otras lenguas en la ontología.

La formalización de otras denominaciones y de equivalentes en inglés se ha realizado a través de una anotación de tipo *label*. Este tipo de anotaciones permite agregar una etiqueta a cualquier elemento de la ontología y asignarle un código de idioma. Por ejemplo, en la figura 127, se han remarcado las etiquetas de tipo *label* de la clase ‘aspecto superficial’. Al lado del contenido de cada etiqueta, se observa la definición del idioma de la etiqueta, @en para inglés, @es para español.

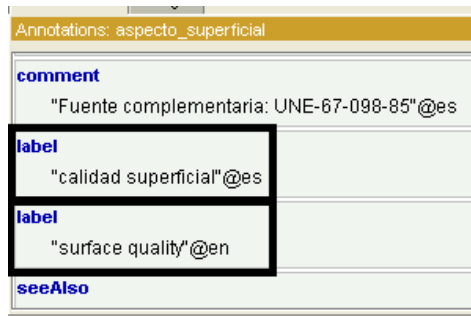


Figura 127. Anotaciones sobre otra denominación y el equivalente en inglés de ‘aspecto superficial’

Este tipo de anotación, que se puede realizar en cualquier tipo de elemento de la ontología, se tiene en cuenta en las tareas de consulta de la ontología o representación de conceptos. Por ejemplo, podemos configurar la aplicación Protégé para ver la ontología en una determinada lengua de manera que todos los elementos que tengan una etiqueta *label* en la lengua escogida se visualicen en esa lengua. En la siguiente figura mostramos un fragmento estructura de clases de ‘característica de la baldosa cerámica’ a través de sus denominaciones en inglés.

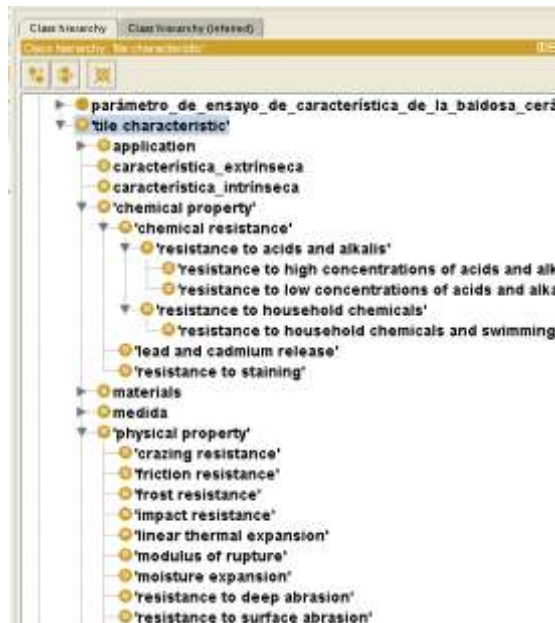


Figura 128. Estructura de clases vista a través de las denominaciones en inglés

Igualmente, podemos hacer consultas, sin importar en que lengua estamos visualizando la ontología, utilizando denominaciones en una lengua diferente de la que se ha utilizado para crear la ontología de conceptos. Por ejemplo, podemos realizar una consulta en la pestaña *DL Query* sobre un tipo de baldosa y escribir esta consulta usando las denominaciones en inglés.



Figura 129. Consulta sobre las clases subordinadas a la clase 'resistencia química' a través de su denominación en inglés

6.10 Resumen de los elementos que componen la ontología

A continuación, presentamos un resumen de cómo hemos formalizado cada uno de los elementos que conforman la ontología a través de su posición en las distintas pestañas de la ontología vista desde la herramienta Protégé.

Las características, actividades y entidades de la ontología se han formalizado mediante clases y se han estructurado en forma de clases y subclases basándose en la información obtenida durante la fase de extracción de información. En la figura 130, se visualiza un fragmento de la jerarquía de clases de la ontología, la jerarquía de clases incluye características de la baldosa, parámetros de ensayo de la característica, instrumentos de medida de las características de la baldosa, tipos de baldosa cerámica (comerciales y según las normas), procesos de ensayo de la baldosa y procesos de fabricación de la baldosa. Las subclases de 'característica', 'actividad' y 'entidad', al pertenecer al ámbito de la cerámica, se han formalizado también como subclase de 'cerámica'.

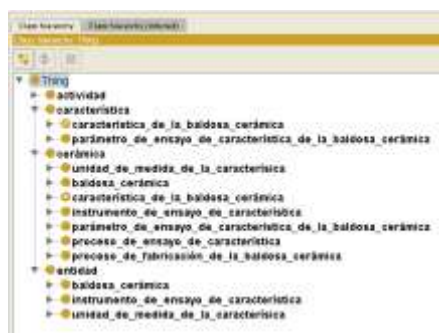


Figura 130. Fragmento de la jerarquía de clases de la ontología

Las características se han formalizado a través de clases y a través de propiedades. Un ejemplo de característica formalizada como clase se muestra en la figura 131, esta clase contiene la siguiente información de izquierda a derecha y de arriba a abajo:

- Clase situada en la jerarquía de clases de 'característica de la baldosa cerámica'
- Contextos declarados en forma de anotación de tipo *comment*
- Referencia a una fuente complementaria de la característica 'aspecto superficial' en forma de anotación de tipo *comment*

- Otra denominación sinónima de ‘aspecto superficial’ en castellano, formalizada a través de una anotación con la etiqueta *label* y con la descripción del idioma (@es)
- Equivalente en inglés de ‘aspecto superficial’, formalizado a través de una anotación con la etiqueta *label* y con la descripción del idioma (@en)
- Referencia a fuente de la característica, anotada a través de la etiqueta *SeeAlso*
- Característica distintiva de la clase ‘aspecto superficial’, declarada a través de la combinación de expresiones OWL 2 en *Equivalent classes* y en *Superclasses*
- Superclase de la clase ‘aspecto superficial’
- Característica esencial de la clase ‘aspecto superficial’, declarada a través de la combinación de expresiones OWL 2 en *Superclasses*

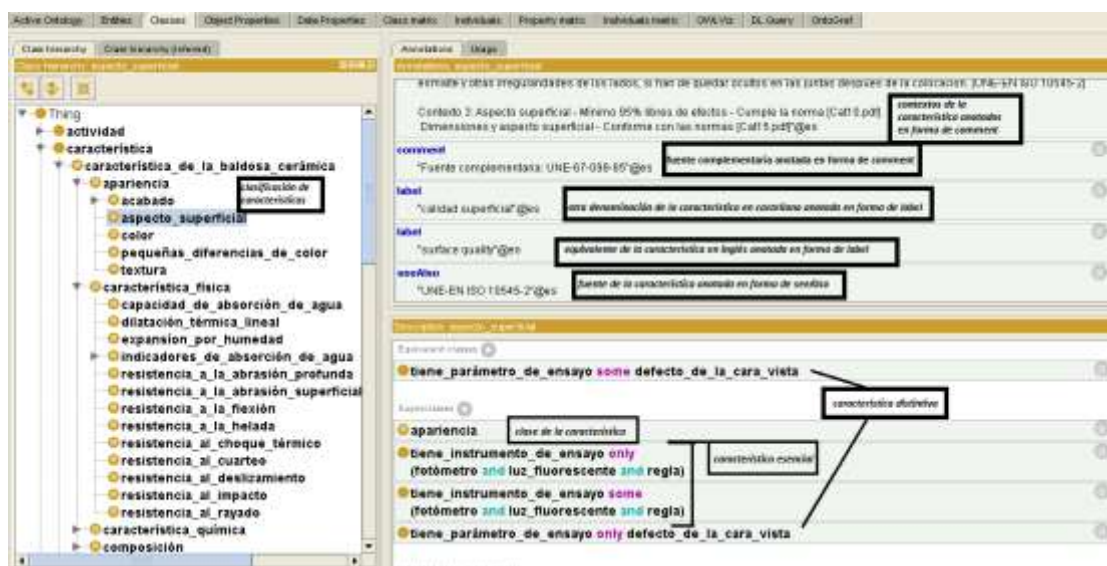


Figura 131. Formalización de la característica ‘aspecto superficial’ desde la pestaña de *Classes*

Además, una clase de característica puede contener información heredada de su superclase, visible en el apartado *inherited anonymous classes*. En el caso de ‘aspecto superficial’, la clase hereda la siguiente información de la superclase ‘apariciencia’ y la superclase ‘característica de la baldosa cerámica’ (figura 132):

- Característica esencial para pertenecer a la clase ‘característica de la baldosa’, declarada mediante la combinación de dos expresiones OWL 2
- Restricción de posibles superclases de ‘aspecto superficial’ heredadas de ‘característica de la baldosa’
- Descripción del tipo de característica heredado de la clase ‘apariciencia’
- Característica esencial para pertenecer a la clase ‘apariciencia’, declarada mediante la combinación de dos expresiones OWL 2

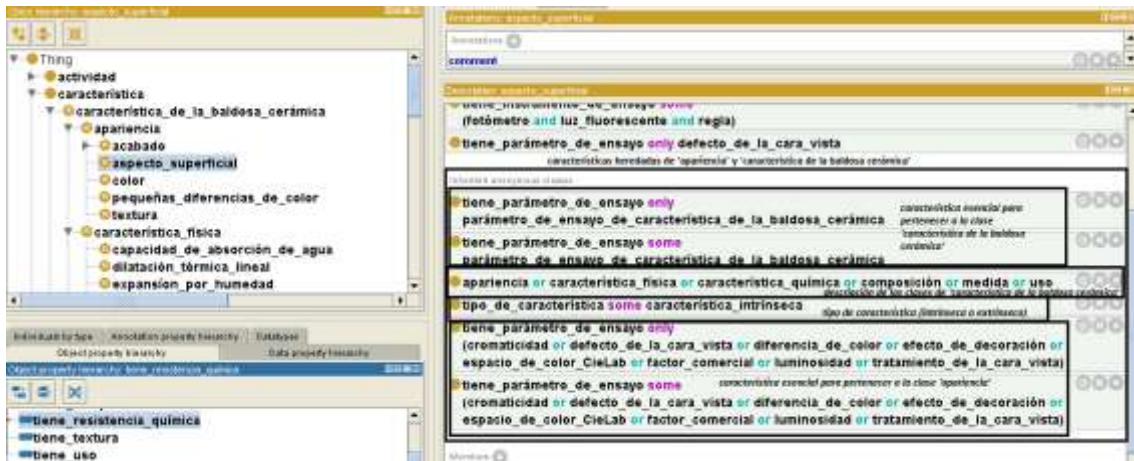


Figura 132. Descripciones heredadas de la superclases de ‘aspecto superficial’

Las características también se han formalizado a través de propiedades para reflejar su rol descriptor y ser utilizada en la descripción de baldosas cerámicas. Siguiendo con el ejemplo, ‘aspecto superficial’ se ha formalizado mediante una *datatype property* denominada ‘tiene aspecto superficial’ donde se incluye la siguiente información, visible en la pestaña *Data properties* del sistema (figura 133):

- Posición en la jerarquía de propiedades
- Observación sobre el tipo de valor en forma de anotación de tipo *comment*
- Declaración de la clase que define la propiedad en la ontología
- Referencia a la clase que representa el parámetro utilizado para determinar los valores de la propiedad en forma de anotación de tipo *seeAlso*
- Declaración del dominio de la característica
- Declaración del tipo de valor que puede tener la característica (en este caso, es *integer*, un número entero)
- Grupo de propiedades al que pertenece la propiedad

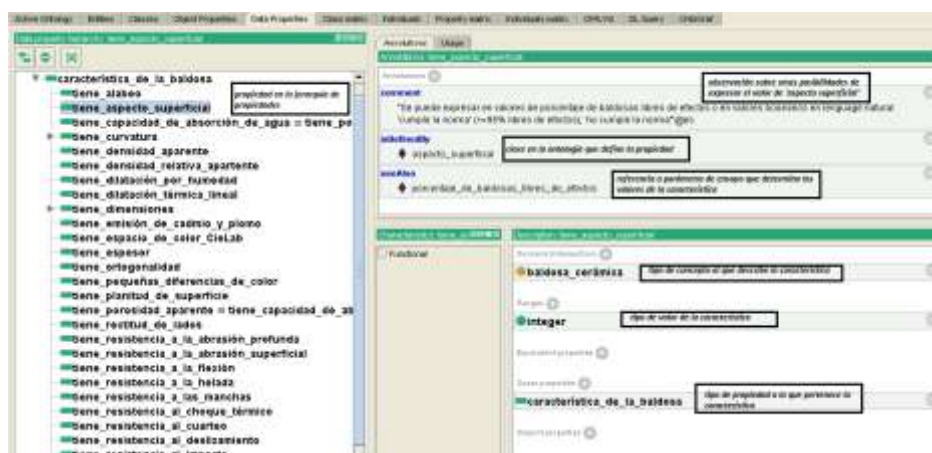


Figura 133. Formalización de la propiedad que representa a ‘aspecto superficial’ desde la pestaña *Data properties*

Las clases que representan a los procesos de ensayo de las características, los instrumentos de ensayo, los parámetros de ensayo, las unidades de medida y los procesos de fabricación de la baldosa no se han descrito a través de características distintivas y esenciales. En algunos casos se ha formalizado alguna observación o información relevante en forma de anotación de tipo *comment*. Por ejemplo, en la figura 134, se observa una descripción del significado de la variable *P* que mide la porosidad aparente.

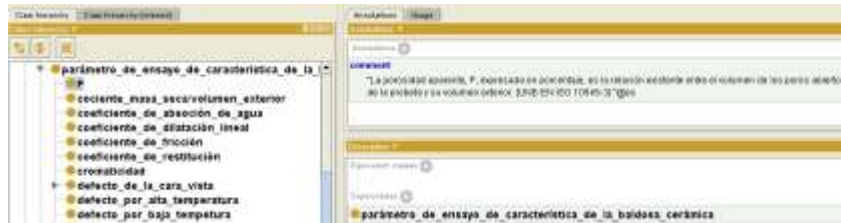


Figura 134. Descripción de la clase 'P' que representa el parámetro de ensayo de 'porosidad aparente'

La relación *característica – proceso de ensayo* y la relación *baldosa – proceso de fabricación* se han formalizado como *object property*, de la misma manera que se ha formalizado el rol descriptor de las características con valores de tipo léxico.

Por ejemplo, en la figura 135, vemos la formalización de la relación característica – proceso de ensayo. En esta relación formalizada a través de una *object property* se ha declarado la siguiente información:

- Tipo de object property
- Referencia a la clase que representa los procesos de ensayo a través de una anotación con la etiqueta *IsDefinedBy*
- Clase que puede participar en la relación, formalizado a través del dominio de la propiedad
- Clase que representa los procesos de ensayo de la relación, formalizado a través del rango de la propiedad.

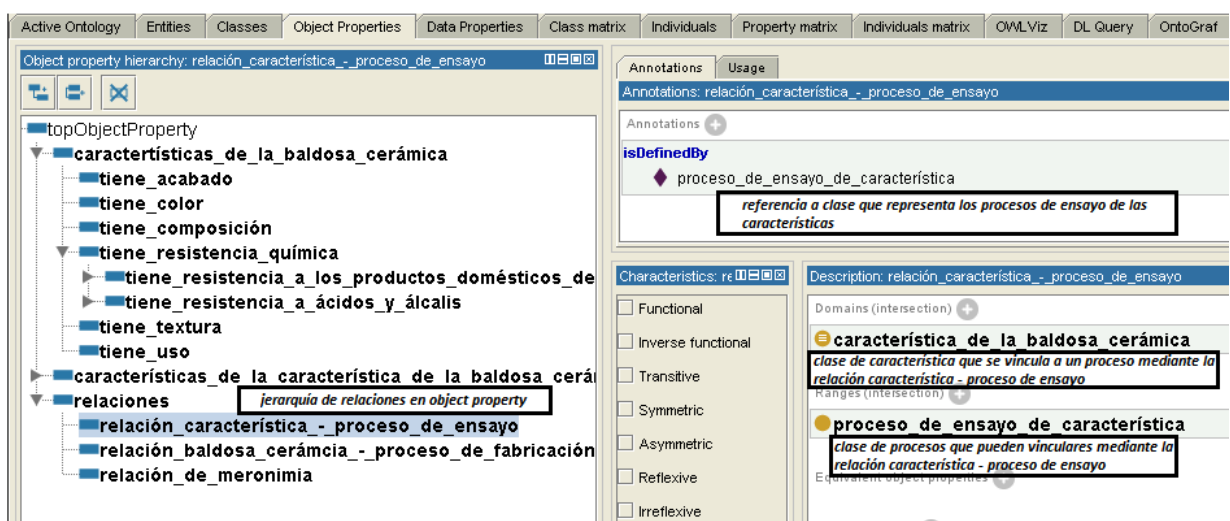


Figura 135. Formalización de la relación *característica – proceso de ensayo*

La relación de meronimia incluye además una declaración en la clase que constituye el todo y en las clases que son parte, como se describe en III.6.5.3.

La descripción de las clases de baldosa cerámica puede incluir información sobre su clase o subclase y las características esenciales y distintivas o la relación con algún proceso de fabricación. En la descripción de las características distintivas o la relación distintiva de la baldosa cerámica, los formalismos coinciden con los utilizados para describir las características distintivas en las características de la baldosa cerámica pero difieren en el criterio: las hermanas de una clase de baldosa cerámica no tienen porqué distinguirse entre ellas a partir una misma característica compartida con valor diferente.

La descripción de las características esenciales de las baldosas o la relación baldosa – proceso de fabricación difiere con respecto a las características de la baldosa en la delimitación de las características o la relación, al permitir que una clase de baldosa agrupe a baldosas que no tengan todas las características declaradas como esenciales pero no incumplan con ninguna de las restricciones sobre características.

7 Introducción de datos en la ontología

Una vez probada y analizada la forma en que vamos a representar nuestros conceptos en la ontología, hemos llevado a cabo la implementación de ésta. A continuación, explicamos la metodología que hemos utilizado y los resultados obtenidos.

7.1 Creación de clases y estructura de clases

El primer paso ha sido crear las clases y subclases que representan los distintos conceptos de la ontología. Hemos creado las clases: ‘cerámica’, ‘entidad’, ‘actividad’ y ‘característica’. Posteriormente, hemos creado las subclases de la clase ‘entidad’, ‘actividad’ y ‘característica’ dando lugar a la estructura representada en la figura 136.

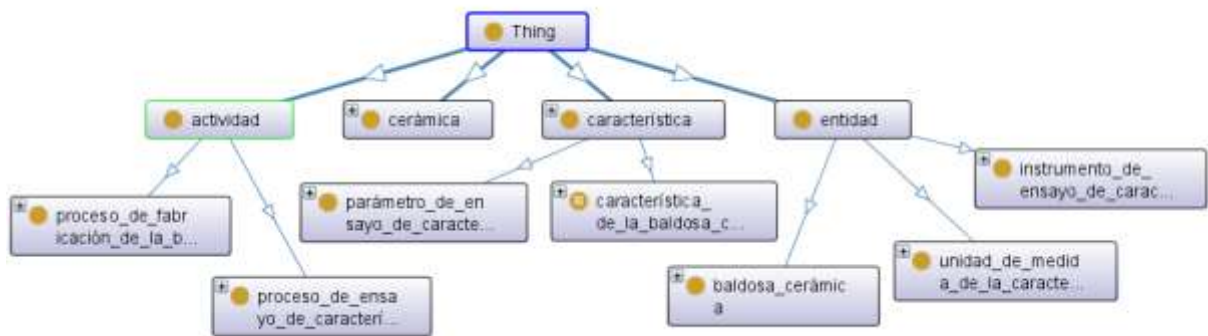


Figura 136. Primer nivel de clases de la ontología

A continuación, hemos relacionado las clases vinculadas a la cerámica con la clase ‘cerámica’. De esta forma una misma clase, por ejemplo ‘instrumento de medida de característica’, pertenece a la clase ‘entidad’ y al mismo tiempo a la clase ‘cerámica’ y de esta forma, se permite la reutilización de la ontología en conjunto con otros dominios (por ejemplo, agregando otros tipos de entidad que no sean de la cerámica en la ontología) sin perder expresividad (al declarar, por ejemplo, que los miembros de la clase ‘instrumento de medida de característica’ en este contexto pertenecen exclusivamente al ámbito de la cerámica). Esta formalización da lugar a una doble clasificación o doble herencia donde una misma clase tendrá dos superclases (figura 137).

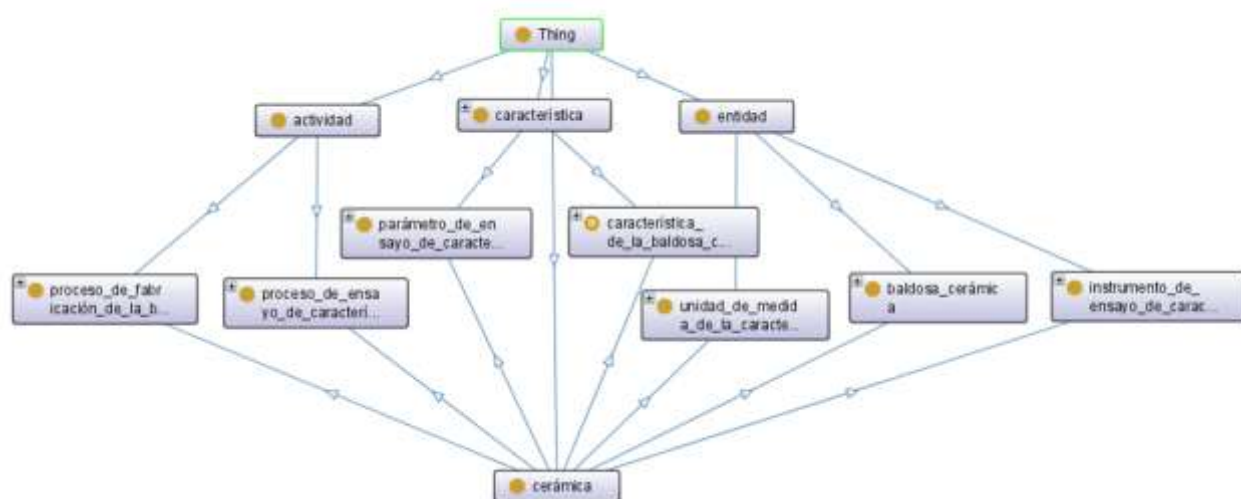


Figura 137. Visualización de la herencia múltiple en clases de la cerámica

El siguiente paso ha sido la creación de subclases para las clases de cada último nivel (característica de la baldosa, defecto de la baldosa, etc.), y a continuación, hemos realizado el mismo paso para los nuevos últimos niveles hasta completar la estructura de clases de la ontología. Por ejemplo, en el anexo 9 puede consultarse el detalle del grafo de la clase 'característica de la baldosa cerámica'.

Para introducir el resto de subclases de cada concepto hemos utilizado la herramienta de Protégé *Create Class Hierarchy*. Primero, hemos representado en distintos documentos Word las estructuras de clases y subclases que queríamos introducir según cada gran grupo, la estructura se ha jerarquizado mediante tabuladores. Una vez recogidas todas las estructuras, hemos ido introduciéndolas una a una como subclases de cada gran grupo. En total, hemos elaborado los siguientes documentos:

- estructura de baldosas cerámicas (anexo...)
- estructura de características
- estructura de instrumentos
- estructura de parámetros
- estructura de actividades

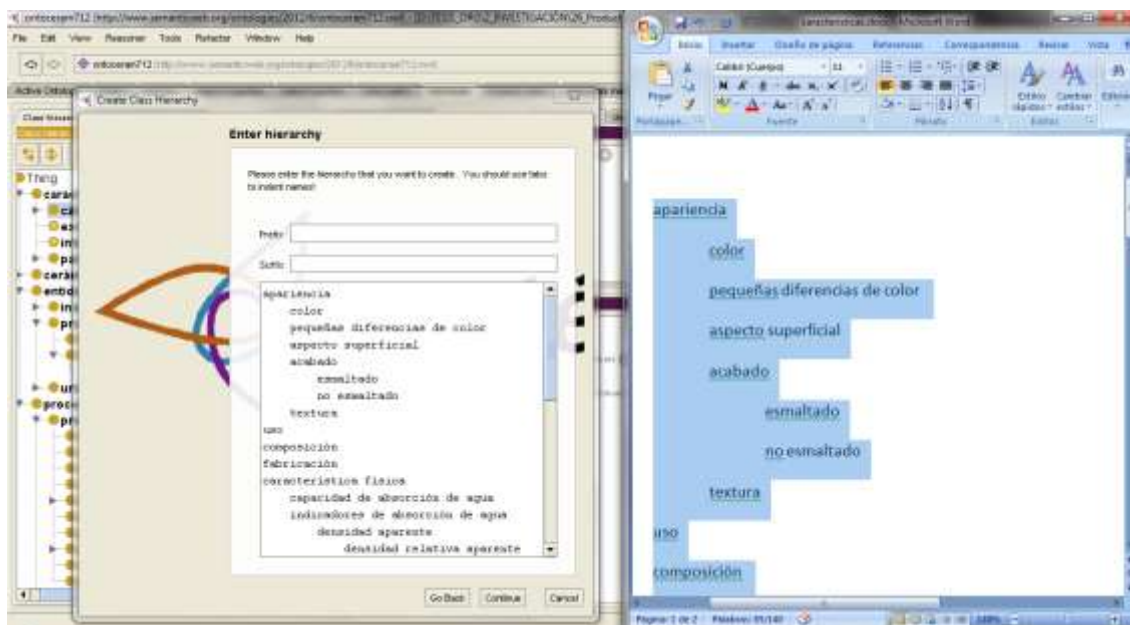


Figura 138. Creación de clases y subclases desde *Create Class Hierarchy* en Protégé y ejemplo de estructura desde Word

Una vez representadas todas las clases de la ontología, se han declarado las clases disjuntas entre clases de la ontología. Concretamente, se han declarado como clases disjuntas entre ellas:

- ‘actividad’, ‘entidad’ y ‘característica’
- Las subclases de ‘actividad’ (‘proceso de ensayo de característica’ y ‘proceso de fabricación de la baldosa cerámica’)
- Las subclases de ‘entidad’ (‘baldosa cerámica’, ‘instrumento de ensayo de características’ y ‘unidad de medida’)

En el caso de las subclases de ‘característica’, ‘característica de la baldosa cerámica’ y ‘parámetro de ensayo de características’ no se han formalizado como disjuntas porque consideramos que pueden compartir miembros. En la recogida de datos, hemos encontrado un caso donde una característica de la baldosa tiene como parámetro otra característica que a su vez es característica de la baldosa, es el caso de ‘capacidad de absorción de agua’ que tiene como parámetro la clase ‘indicadores de absorción de agua’. Esta clase agrupa las características de la baldosa cerámica ‘porosidad aparente’, ‘densidad aparente’ y ‘densidad relativa aparente’. Además, ‘porosidad aparente’ es equivalente a ‘capacidad de absorción de agua’, aspecto que no resulta incompatible con lo anterior (figura 139).

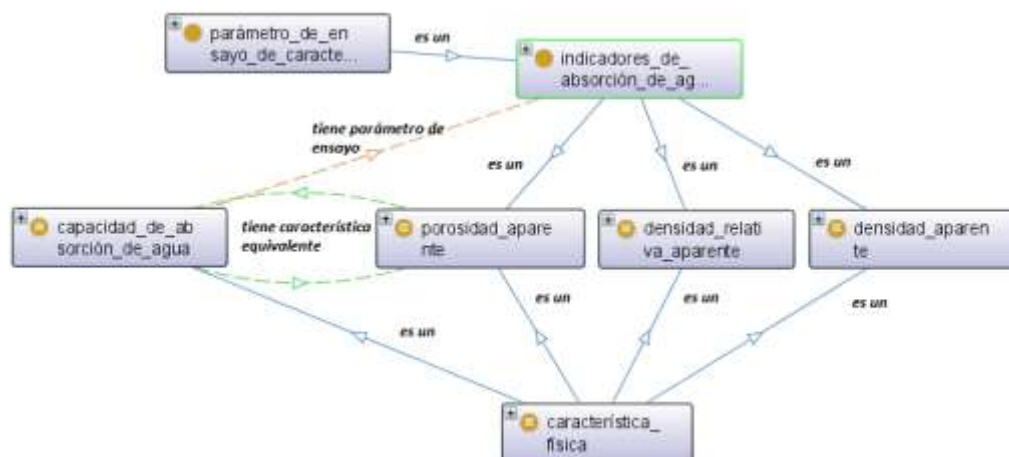


Figura 139. Grafo sobre 'indicadores de absorción de agua' y relaciones entre conceptos subordinados

7.2 Introducción de características en forma de *object property* e introducción de valores

El siguiente paso de introducción de datos ha sido la creación de *Object property* para las características que se han formalizado con valores de tipo léxico y que se representarán como clases en la ontología. Esta tarea se ha desarrollado desde la ventana *Classes* y la ventana *Object Properties*.

La mayoría de características recogidas tienen valores numéricos, booleanos u otro tipo de valor que se recogerá en forma de *Datatype Property*, por lo que la lista de características de tipo *Object Property* se limita a:

1. acabado
2. color
3. composición
4. fabricación
5. resistencia química
6. resistencia a ácidos y álcalis
7. resistencia a ácidos y álcalis en concentraciones débiles
8. resistencia a ácidos y álcalis en concentraciones fuertes
9. resistencia a los productos domésticos de limpieza
10. resistencia a los productos domésticos de limpieza y aditivos para agua de piscinas
11. textura
12. uso

Antes de crear la propiedad que servirá para representar el rol descriptor de la característica en cuestión, debemos introducir los posibles valores de esa característica. Los valores de tipo unidad léxica se han formalizado a través de instancias. Por ejemplo, para la característica 'textura' hemos creado una instancia para cada valor y las hemos declarado como tipo de la clase 'textura'.

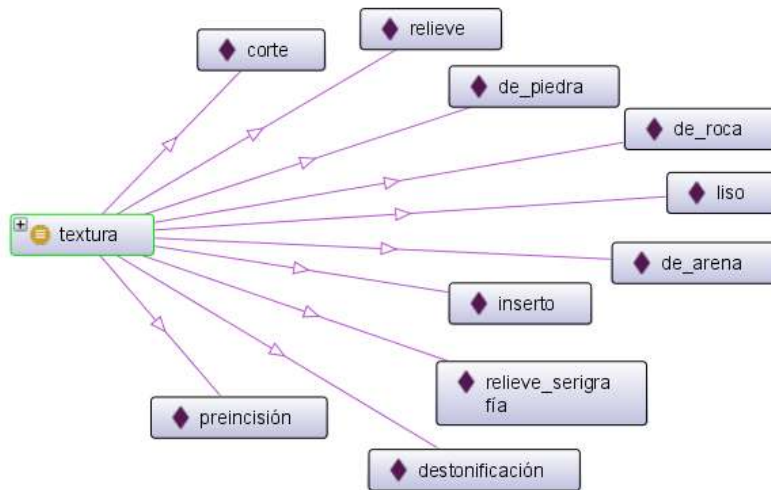


Figura 140. Valores de textura

En el caso de que tuviéramos información u observaciones adicionales que agregar en lenguaje natural, en este paso, hemos agregado esa información a través de una anotación en forma de comentario (subelemento *comment*).

Por ejemplo, hemos agregado la siguiente información sobre el valor 'GA': « Baldosa esmaltada sin efecto visible [UNE-EN 10545-13]» que ayuda a explicar el significado del valor.

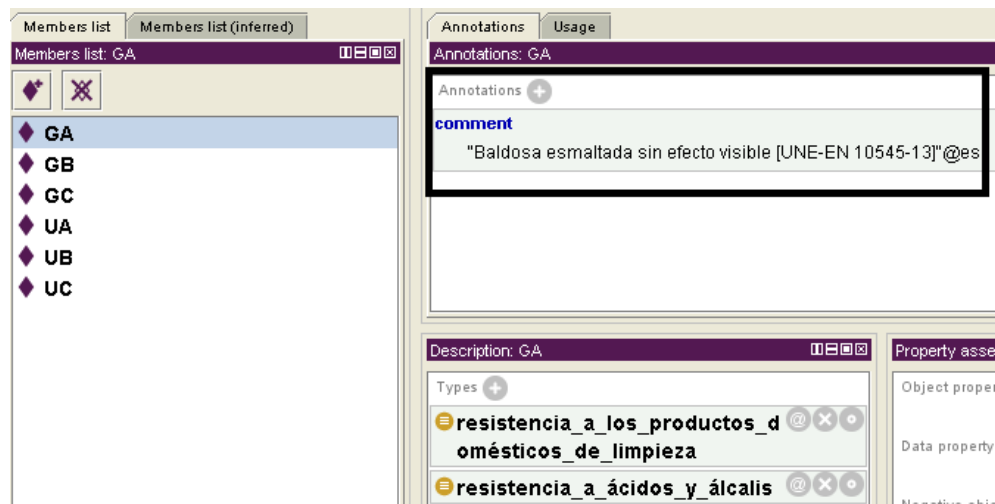


Figura 141. Descripción en lenguaje natural del valor pasta roja en la ontología

Este proceso se ha repetido para cada característica de la lista anterior.

El siguiente paso ha sido la creación de una *Object Property* para cada característica y la asignación del dominio y el rango. La estructura de las *object property* distingue entre 'características del baldosa cerámica', 'características de las características de la baldosa cerámica' y 'relaciones'. Las propiedades se han creado bajo un tipo u otro según su naturaleza.

Las características de la baldosa representadas en forma de *Object Property* de nuestra ontología tiene como dominio 'baldosa cerámica', 'baldosa esmaltada' o 'baldosa no esmaltada' según la característica y el rango viene determinado por la característica concreta. En el caso de la propiedad 'tiene acabado' es 'acabado', en el caso de 'tiene composición' es 'composición' y

así para cada clase de característica respectivamente. Además al crear la *object property* para una característica, hemos indicado si la característica es funcional o no, marcando la casilla de *functional* o dejándola desmarcada. A continuación, mostramos un ejemplo de la *object property* creada para ‘acabado’.



Figura 142. Descripción en forma de Object property de ‘tiene acabado’

En esta fase, también hemos introducido una referencia al parámetro utilizado para asignar los valores de la propiedad. Esta tarea se ha realizado desde la ventana de anotaciones, utilizando el subelemento *seeAlso*. Por ejemplo, en ‘tiene acabado’ se ha creado una anotación que hace referencia al parámetro de ensayo ‘tratamiento de la cara vista’:

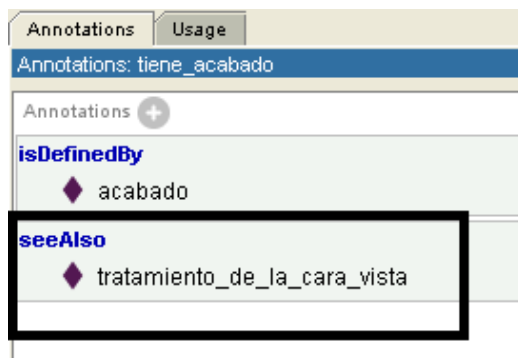


Figura 143. Referencia al parámetro de ensayo de ‘tiene acabado’

Finalmente, el último paso en esta fase de trabajo es enlazar la propiedad creada con su clase en la ontología. Para ello, hemos creado una anotación y hemos utilizado el subelemento *IsDefinedBy* que sirve para vincular elementos de la ontología con otros elementos (figura 144).

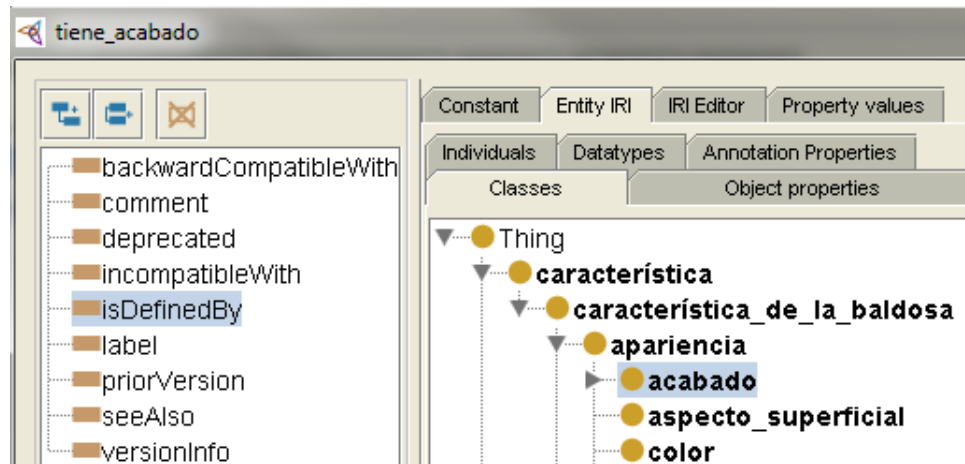


Figura 144. Anotación sobre el vínculo entre la clase ‘acabado’ y la propiedad ‘tiene acabado’

En resumen, la introducción del rol descriptor de una característica con valor léxico en la ontología pasa por:

- a. Introducir los valores de la característica en forma de instancias de su clase.
- b. Agregar comentarios en forma de anotación sobre los valores, si es pertinente.
- c. Crear una ObjectProperty para la característica indicando el dominio (clase que contiene entidades susceptibles de ser descritas por la característica) y el rango (clase que contiene los valores posibles que puede tomar la característica).
- d. Enlazar la propiedad con su clase en la ontología (IsDefinedBy).

7.2.1 Casos particulares

Algunas características de este tipo han conllevado además otras consideraciones a la hora de llevar a cabo la introducción de datos. Por ejemplo, la característica ‘acabado’ distingue, como corroboró el experto, entre valores de esmaltado y no esmaltado, y dentro del valor ‘esmaltado’ puede concretarse en ‘brillo’, ‘mate’, ‘satinado’, ‘nácar’, ‘rejunte’ y ‘metalizado’. Igualmente el valor de no esmaltado se puede concretar entre ‘pulido’ y ‘sin pulir’. Para introducir esta especificación hemos creado primero dos subclases de ‘acabado’ (‘esmaltado’ y ‘no esmaltado’) y a continuación, hemos agregado los valores correspondientes en forma de instancia (figura 145).

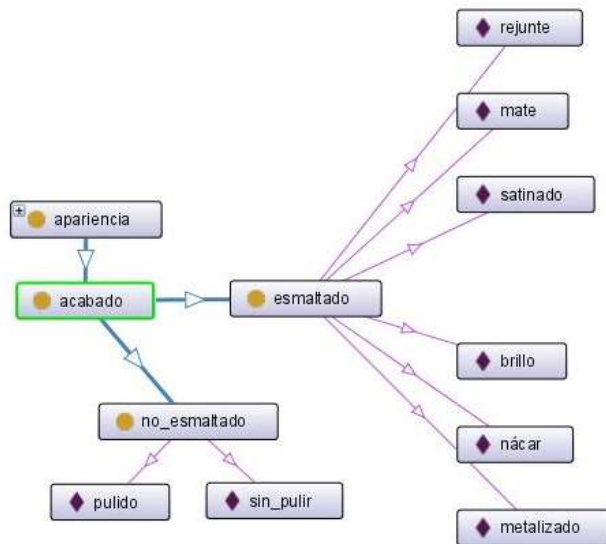


Figura 145. Formalización de los valores de ‘acabado’ en la ontología

Esto mismo ha ocurrido con los valores de ‘uso’, donde hemos realizado una primera distinción entre ‘suelo’ y ‘pared’, y con los valores de ‘composición’, donde hemos realizado una primera distinción entre ‘pasta blanca’ y ‘pasta roja’.

Otro caso particular es el ‘color’, donde además de la creación de la *ObjectProperty*, hemos facilitado la posibilidad de vincular los colores con las coordenadas en el espacio de color CieLab. Para ello, hemos creado una *DatatypeProperty* denominada ‘tiene_espacio de color CieLab’ que podrá utilizarse tanto para describir los valores CieLab de un color, como el color de una baldosa directamente. (figura 146).

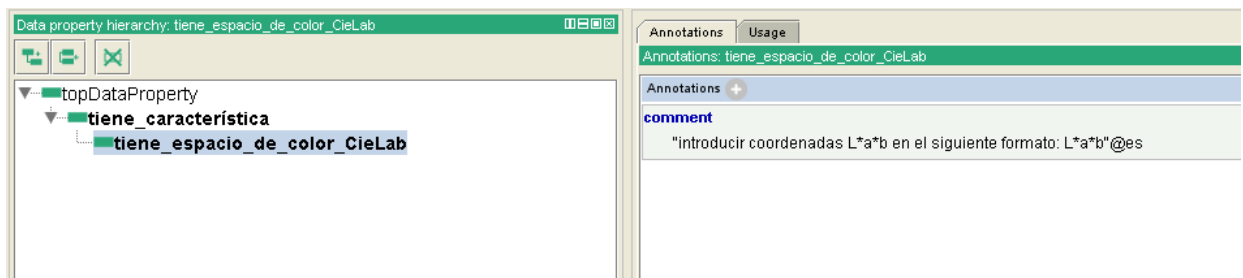


Figura 146. Complemento para la descripción de color

Si se quisiera profundizar en este aspecto, se podría además agregar una *sub-Data property* para cada coordenada L^*a^*b . En este trabajo no hemos profundizado en este aspecto.

7.3 Introducción de características en forma de *datatype Property* e introducción de valores

El rol descriptor de las características que tienen valores numéricos (limitados o ilimitados) o booleanos se ha formalizado a través de *datatype property* que enlaza un concepto de la ontología con un tipo de valor o con un valor concreto.

La lista de características de este tipo en la ontología es extensa y comprende:

1. alabeo
2. anchura
3. aspecto superficial
4. capacidad de absorción de agua
5. curvatura
6. curvatura central
7. curvatura lateral
8. densidad aparente
9. densidad relativa aparente
10. dilatación por humedad
11. dilatación térmica lineal
12. dimensiones
13. emisión de cadmio y plomo
14. espesor
15. longitud
16. ortogonalidad
17. pequeñas diferencias de color
18. planitud de superficie
19. porosidad aparente
20. rectitud de lados
21. resistencia a la abrasión profunda
22. resistencia a la abrasión superficial
23. resistencia a la flexión
24. resistencia a la helada
25. resistencia a las manchas
26. resistencia al choque térmico
27. resistencia al cuarteo
28. resistencia al deslizamiento
29. resistencia al impacto
30. resistencia al rayado

De entre estas características, las que toman valores booleanos son:

1. aspecto superficial
2. pequeñas diferencias de color
3. resistencia a la helada
4. resistencia al choque térmico
5. resistencia al cuarteo
6. resistencia al impacto

El resto, toman valores numéricos limitados o ilimitados, y hacen referencia en la mayoría de los casos a unidades de medida que aparecen relacionados en la ontología.

La creación de *Datatype* para representar las características con valores booleanos se realiza en la pestaña *Data property*. El dominio de la característica es el producto acabado cerámico y como valor hemos indicado que se trata de un valor de tipo *boolean* (figura 147).

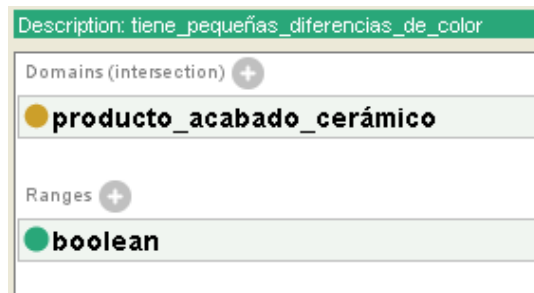


Figura 147. Formalización del dominio y rango de ‘pequeñas diferencias de color’

En el caso de disponer de expresiones en lenguaje natural para indicar el valor positivo o negativo que toma la característica, hemos incluido esta información en forma de anotación. Por ejemplo, en el caso de ‘tiene resistencia a la helada’ las referencias en los recursos textuales indican que se dice “pasa la prueba” o “pasa el ensayo” en caso de cumplir con lo establecido para tener la característica y “no pasa la prueba” o “no pasa el ensayo” en caso negativo (figura 148).

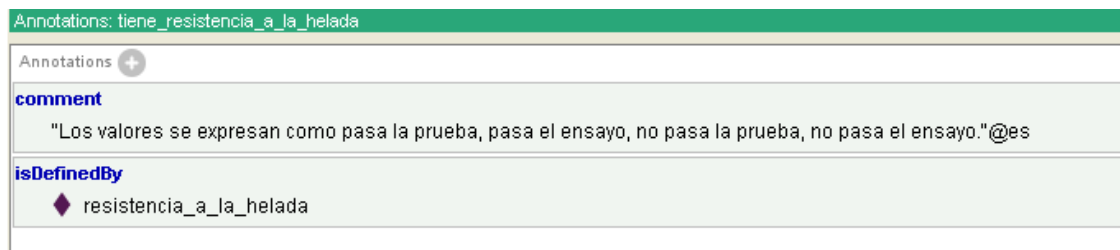


Figura 148. Anotaciones en ‘tiene resistencia a la helada’

Igualmente, en todas las características presentadas en este apartado, al igual que en el caso de las *object property*, hemos enlazado la propiedad con la clase en cuestión que describe la característica a través del elemento “IsDefinedBy” (figura 148).

Las características que tienen como valores datos numéricos se han creado siguiendo un proceso similar al anterior. El aspecto que cambia es la atribución del tipo de valor y la inclusión de la relación entre la propiedad y la unidad de medida utilizada en caso necesario.

El dominio de las propiedades que representan a este tipo de características es el ‘la baldosa cerámica’, excepto: ‘emisión de cadmio y plomo’, ‘resistencia a la abrasión superficial’ y ‘resistencia al rayado’ que tienen como dominio ‘baldosa esmaltada’ y ‘resistencia a la abrasión profunda’ cuyo dominio es exclusivamente ‘baldosa no esmaltada’.

El rango variará según la característica pudiendo ser de tipo *double* (para valores numéricos con decimales), *integer* (para enteros) o *string* (para ‘dilatación térmica lineal’) y además, pudiendo estar delimitado o ser ilimitado según la característica en concreto. Veamos unos ejemplos de cómo hemos introducido esta información.

La característica ‘alabeo’ tiene como valores números decimales acompañados de una unidad de medida (mm). Para introducir esta información en la ontología, primero, hemos creado una *datatype property* denominada ‘tiene_alabeo’ cuyo dominio es ‘producto_acabado_cerámico’ y cuyo rango es un valor de tipo *double*.

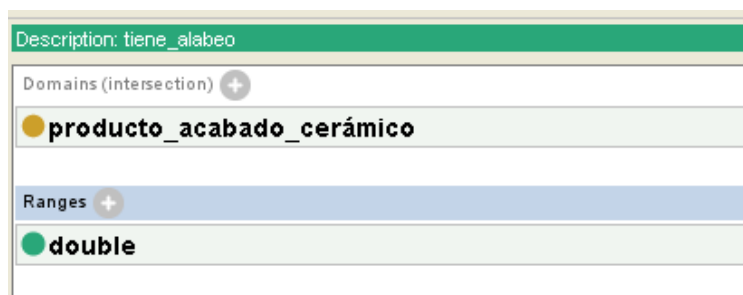


Figura 149. Dominio y rango de ‘tiene alabeo’

En el caso de características ‘resistencia a la abrasión superficial’, ‘resistencia a las manchas’, ‘resistencia al deslizamiento’ y ‘resistencia al rayado’ cuyo rango de valores es restringido, la descripción del rango varía con respecto a lo explicado en el párrafo anterior. La creación de la propiedad se ha realizado de la misma manera, hemos creado la *datatype property* y le hemos asignado un dominio. Por ejemplo, ‘resistencia a la abrasión superficial’ tiene como dominio ‘baldosa esmaltada’. En el caso de rango, hemos introducido el tipo de valor que puede tener (double o integer) y seguido la delimitación de los valores, formando una expresión en OWL 2. A continuación, se muestra la delimitación de los valores de ‘tiene resistencia a la abrasión superficial’ (figura 150).

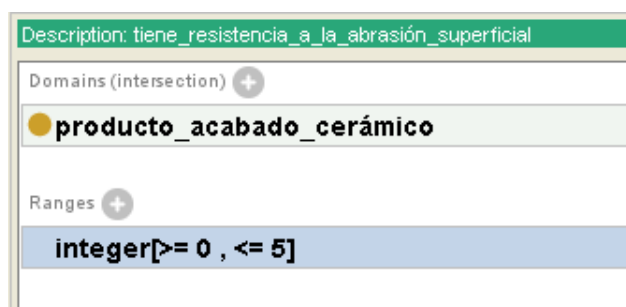


Figura 150. Restricción de valores posibles en “tiene a la abrasión superficial”

El resto de la descripción se ha realizado siguiendo los mismos pasos que con el resto de propiedades. Finalmente, en el caso de ‘dilatación térmica lineal’, como hemos comentado en §6.3.7, el valor asignado para la propiedad ha sido *string*.

Las características que aluden al parámetro o a la unidad de medida de una característica de la baldosa cerámica se han representado en las clases de características de la baldosa. No obstante, también hemos introducido una anotación a través del subelemento *seeAlso* sobre la unidad de medida en la que se expresa el valor de la característica o el parámetro que debe considerarse. Por ejemplo, en el caso de ‘alabeo’ existe una referencia a la unidad de medida ‘milímetro’.



Figura 151. Anotaciones de 'tiene_alabeo'

En resumen, la introducción del rol descriptor de una característica con valor léxico en la ontología pasa por:

- a. Introducir los valores de la característica en forma de instancia
- b. Crear una *datatype property* para la característica indicando el dominio (clase que contiene entidades susceptibles de ser descritas por la característica) y el rango (clase que contiene los valores posibles que puede tomar la característica).
- c. Limitar el rango en caso necesario.
- d. Agregar comentarios en forma de anotación, si es pertinente.
- e. Enlazar la propiedad con parámetros y/o unidades de medida relacionados con la característica, si es pertinente.
- f. Enlazar la propiedad con su clase en la ontología.

Casos particulares

Un caso particular en la introducción de datos ha sido el de la característica 'dilatación por humedad' que se puede expresar bien en milímetro por metro o bien en porcentaje. Esto se ha reflejado en forma de anotación indicando que la medida a seguir es el milímetro por metro pero que existe la posibilidad de expresar el valor en porcentaje de expansión (figura 152).

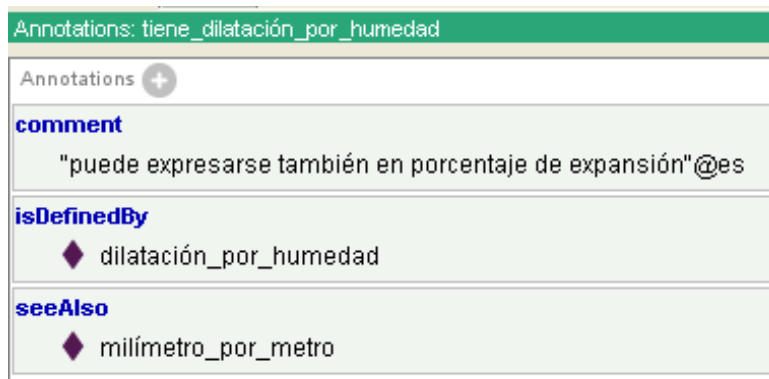


Figura 152. Anotaciones de 'tiene dilatación por humedad'

Otro caso particular similar es el de 'aspecto superficial', en este caso la característica puede expresarse mediante un valor booleano (cumple la norma, no cumple la norma) o un porcentaje de baldosas libres de efectos. Como en el caso anterior, hemos declarado como valor posible un

único tipo de valor, un número entero que expresa el porcentaje de baldosas libres de efectos, y hemos introducido un comentario aludiendo a la posibilidad de utilizar otro tipo de valor para la misma característica.

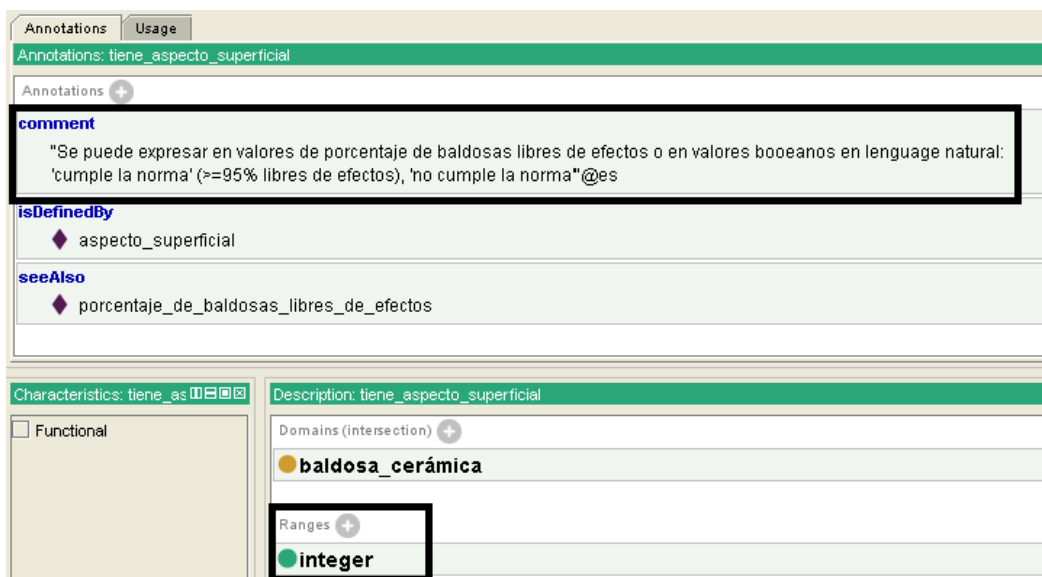


Figura 153. Dominio y rango de 'aspecto superficial'

7.4 Creación de características y relaciones para describir las características de la baldosa

El siguiente paso en la introducción de datos ha sido la creación de una *object property* para cada una de las características y relaciones que describen las características de la baldosa. En el proceso de creación se ha asignado un rango y dominio a cada *object property*.

Las propiedades para representar el tipo de característica, la característica equivalente, el parámetro de ensayo, el instrumento de medida y la unidad de medida se han creado bajo la propiedad 'características de la característica de la baldosa cerámica'. La propiedad que representa la relación de meronimia y la relación característica – proceso de ensayo se ha creado bajo la propiedad 'relaciones'.

Una vez creadas, se ha asignado un dominio y un rango a cada propiedad. El dominio de todas las propiedades es 'característica de la baldosa cerámica' y el rango depende de la propiedad en concreto, por ejemplo, en el caso de 'tiene parámetro' es 'parámetro de ensayo de característica de la baldosa cerámica'.

7.5 Descripción de las clases de característica

Una vez se ha creado la estructura de clases de características y el rol descriptor de cada característica está formalizado a través de las propiedades, incluyendo la clase de baldosas que describe y los valores que puede tomar, en esta fase de trabajo nos hemos centrado en agregar el resto de información recogida sobre la característica:

- Tipo de característica
- Características y relaciones de la característica
- Fuente de la característica y fuentes complementarias

- Contextos y otras observaciones
- Descripción de datos introducidos en lenguaje natural

7.5.1 Tipo de característica

Como se explica en 6.8, para indicar que la característica es extrínseca o intrínseca hemos creado primero dos clases de ‘característica’ que son ‘extrínseca’ e ‘intrínseca’. Además, hemos creado una *object property* denominada ‘tipo de característica’ cuyo dominio es ‘característica de la baldosa’ y cuyo rango es ‘extrínseca’ o ‘intrínseca’ (figura 154).



Figura 154. Definición de la *object property* ‘tipo de característica’

A continuación, hemos descrito en cada clase de característica el tipo de característica que es usando las expresiones “tipo de característica *some* intrínseca” o “tipo de característica *some* extrínseca”, según la característica, que indican que la clase tiene la propiedad ‘tipo de característica’ y que su rango es vinculado a la clase ‘intrínseca’ o ‘extrínseca’, según el valor introducido para cada clase. Dado que las clases de característica y sus subclases coinciden en ser intrínsecas o extrínsecas, la descripción se realiza en las clases de primer nivel de las características de la baldosa. En la figura 155, se visualiza la descripción de la característica ‘apariencia’ como intrínseca.

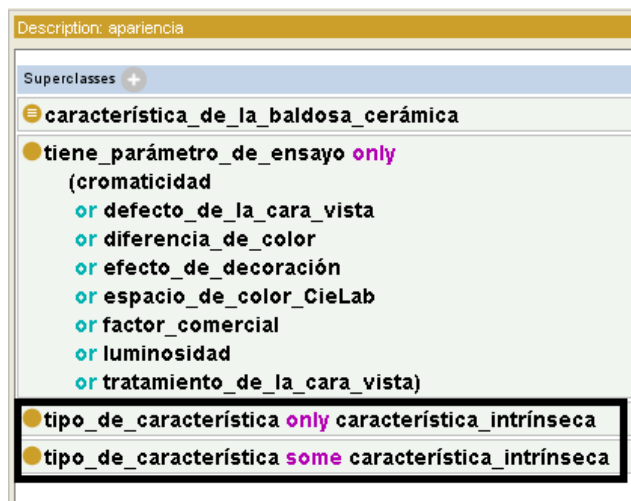


Figura 155. Definición de ‘acabado’ como característica intrínseca

Para la introducción de la información sobre el tipo de característica (intrínseca/extrínseca) y sobre las relaciones entre la característica y otros conceptos de la ontología se ha utilizado la herramienta *Class matrix* que permite agregar los valores de una o

más propiedades a una o más clases de forma rápida gracias a la visualización de los datos en forma de panel (figura 156).



Figura 156. Visualización de extracto de la pestaña *Class matrix* en Protégé 4

En la figura 156, se observa cómo desde el panel de *Class* podemos visualizar todas las clases de la ontología, en el centro hemos seleccionado rellenar los datos para la propiedad ‘tipo de característica’ con la expresión ‘some’. En cada celda, hemos indicado los valores para cada clase de la izquierda. Una de las ventajas de utilizar estas opciones es que una vez asignado un valor a una clase, las subclases toman ese valor, de manera que nos hemos evitado escribir los valores para esas subclases (en la figura, los valores entre paréntesis, se han heredado de su clase).

Finalmente, para agregar contenido a la ontología, en las clases ‘intrínseca’ y ‘extrínseca’, hemos indicado que existe una equivalencia entre ser subclase de ‘intrínseca’ y tener la propiedad ‘tipo de característica’ con valor ‘intrínseca’. Esto mismo hemos realizado con la clase ‘extrínseca’, indicando, en este caso, el valor ‘extrínseca’ (figura 157).



Figura 157. Definición de la clase ‘extrínseca’

De esta forma al pasar el razonador *HermiT* en la ontología la estructura de características cobra una nueva dimensión. Las características de la baldosa aparecen estructuradas como al inicio pero además aparecen como subclases de ‘extrínseca’ o ‘intrínseca’ según el valor tomado.



Figura 158. Estructura inferida por el razonador

7.5.2 Introducción de relaciones, características esenciales y características distintivas

El siguiente paso ha sido la introducción de las relaciones, las características esenciales y las características distintivas de cada característica de la baldosa.

Al tratarse de casos muy poco numerosos, en esta fase, hemos empezado formalizando la relación de meronimia entre ‘planitud de superficie’ y ‘curvatura’ y ‘alabeo’; entre ‘curvatura’ y ‘curvatura central’ y ‘curvatura lateral’; y entre ‘dimensiones’ y ‘espesor’ y ‘ortogonalidad’. Igualmente, hemos formalizado la equivalencia entre ‘capacidad de absorción de agua’ y ‘porosidad aparente’. La formalización de estas clases se ha realizado siguiendo las pautas explicadas en III.6.6.3 y III.6.4, respectivamente donde se describe con detalle la formalización de estos casos concretos.

La relación *característica – proceso de ensayo* se ha introducido en las características correspondientes utilizando el *Class Matrix* (figura 159).

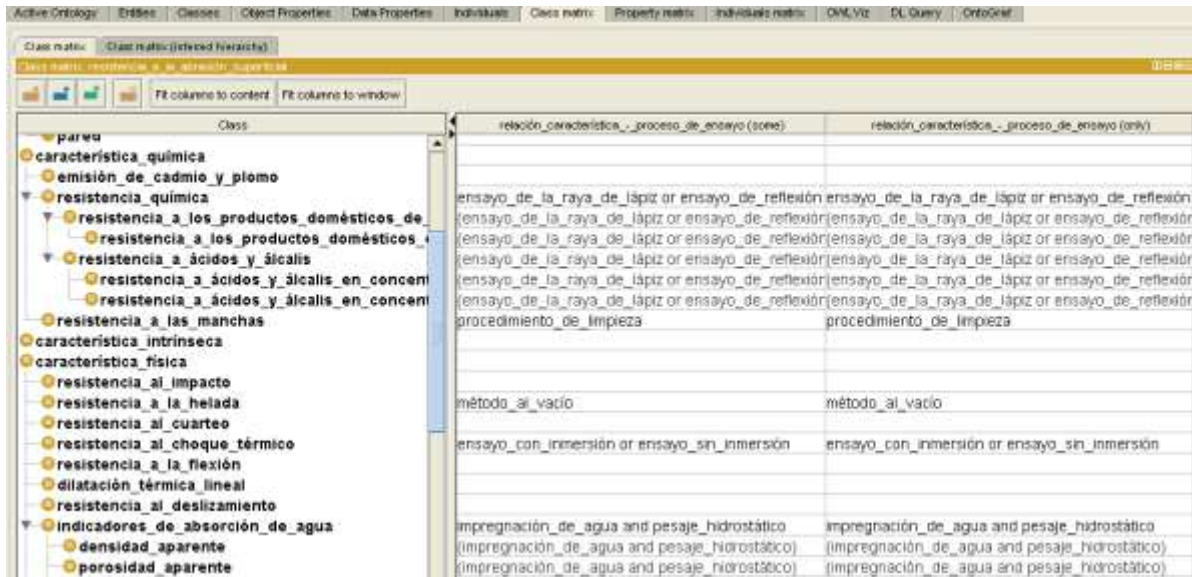


Figura 159. Introducción de la relación *característica – proceso de ensayo* en las características correspondientes

De esta forma, hemos agregado en el apartado *Superclasses* de las clases que correspondía la propiedad *característica – proceso de ensayo* con los valores correspondientes. En total, hemos agregado información sobre el proceso de ensayo de 15 características, concretamente las que se visualizan en la siguiente figura:

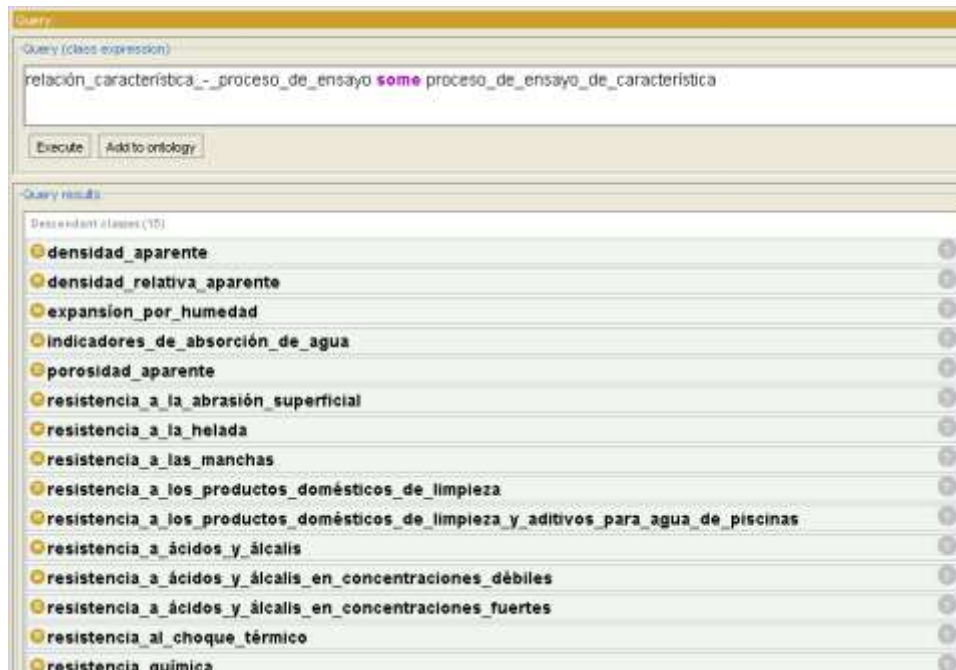


Figura 160. Clases de la ontología que tienen la relación *característica – proceso de ensayo* con algún proceso de ensayo

7.5.2.1 *Introducción de características distintivas y características*

Como hemos explicado en el apartado III.6.8.1, la introducción de las características distintivas y las características esenciales se ha realizado tomando un enfoque *bottom-up*, que se combina con el criterio de diferencia única para identificar la característica distintiva de las

subclases de una clase. La introducción de estas características distintivas se ha realizado de acuerdo con el análisis de los datos obtenidos durante la fase de extracción y representados en la tabla de descripción general de las características.

La introducción de datos se inicia en el último nivel de cada rama, entendiendo por ramas la clase ‘apariencia’, ‘característica física’, ‘característica química’, ‘composición’, ‘medida’ y ‘uso’. El primer paso ha consistido en comparar las características comunes y no comunes entre los miembros de la clase de último nivel, por ejemplo, entre ‘acabado’, ‘aspecto superficial’, ‘color’, ‘pequeñas diferencias de color’ y ‘textura’. La comparación de las características que describen estas clases indica que la característica que todas tienen en común es parámetro. Una vez identificada la característica se han comparado los valores para ver si son distintos entre las clases. En todos los casos, los valores de la característica en común han sido distintos, en parte, debido a que en la mayoría de casos, la característica distintiva es el parámetro de ensayo, cuyos valores aportan un gran nivel de detalle sobre la característica.

El siguiente paso es introducir la característica distintiva y los valores correspondientes a cada clase en el apartado *Equivalent classes* y en el apartado *Superclasses* de cada característica. En este caso no hemos utilizado *Class matrix* sino que la declaración de la expresiones OWL 2 se ha realizado manualmente. El resultado es la formalización de la característica distintiva de cada clase de característica en el nivel.

La característica distintiva más recurrente ha sido ‘tiene parámetro’, esta característica se ha formalizado como distintiva en todas las clases a excepción de:

- Resistencia a los productos domésticos de limpieza
 - Resistencia a los productos domésticos de limpieza y aditivos para agua de piscinas
- Resistencia a ácidos y álcalis
 - Resistencia a ácidos y álcalis en concentraciones débiles
 - Resistencia a ácidos y álcalis en concentraciones fuertes

A continuación, mostramos un fragmento de la descripción de la clase ‘resistencia a ácidos y álcalis en concentraciones fuertes’ con la característica distintiva remarcada en el recuadro (figura 161).



Figura 161. Descripción de la clase 'resistencia a ácidos y álcalis en concentraciones fuertes'

En este caso el parámetro de ensayo que suele ser la característica distintiva más recurrente tiene el mismo valor en esta característica que en su hermana de clase y por tanto, no se puede utilizar como distintiva.

La formalización de las características esenciales de cada característica de la baldosa cerámica se ha realizado comparando el resto de características (no distintivas) de las clases de un nivel para identificar cuáles comparten todas las clases del nivel y cuáles no comparten o toman un valor diferente.

En el caso de que todas las clases de un mismo nivel compartan una característica y su valor, esta característica se ha formalizado en la clase superior (como distintiva o esencial según correspondiera). En el caso de que las clases compartan la característica pero con distinto valor y esta no fuera distintiva, se han formalizado en cada clase del nivel, asignando a cada una el valor correspondiente para la propiedad. Las características no compartidas con otras hermanas de clase también se han formalizado en la clase correspondiente. La formalización se ha realizado siguiendo las pautas descritas en III.6.8.1 para características esenciales.

De esta forma, hemos ido introduciendo todas las descripciones de las clases hasta llegar al nivel de 'característica de la baldosa', donde damos por terminada la formalización. La descripción de esta clase, que se muestra en la figura a continuación, restringe las clases que pueden pertenecer a la clase 'característica de la baldosa cerámica' a través de la expresión "apariciencia or característica_física or característica_química or composición or medida or uso" que indica, por un lado, que para ser clase de 'característica de la baldosa cerámica' una clase debe ser alguna de las enunciadas o subclase de alguna de éstas, y por otro, que la suma de éstas características constituye la extensión de la clase 'característica de la baldosa cerámica'.

Además, se ha declarado que para pertenecer a la clase ‘característica de la baldosa cerámica’, una clase debe contener la propiedad ‘tiene parámetro de ensayo’ con algún valor de la clase ‘parámetro de ensayo de característica de la baldosa cerámica’.

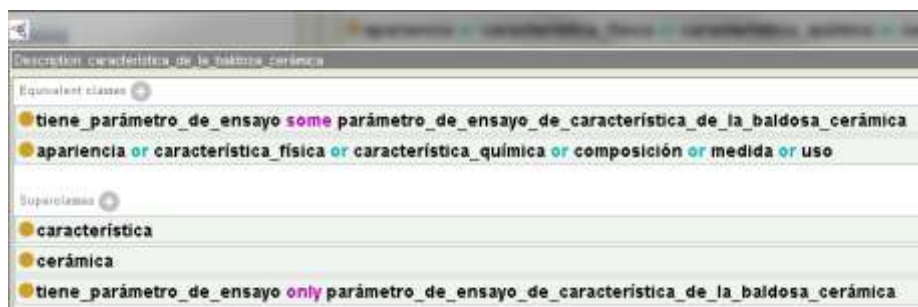


Figura 162. Descripción de la clase ‘característica de la baldosa cerámica’

7.5.3 Contextos

En la pestaña de clases, hemos introducido para cada característica de la baldosa uno o más contextos, según los resultados de la recogida de datos. Estos contextos se han introducido manualmente, copiando y pegando los contextos de la tabla de trabajo a la ontología en Protégé. Concretamente, se han introducido en el apartado ‘comment’ de anotaciones de la clase. La introducción, en cuanto a estructura, incluye la distinción entre “contexto 1” para contextos extraídos del corpus TxtCeram, “contexto 2” para contextos procedentes de las normas y “contexto 3” para contextos procedentes de los catálogos.

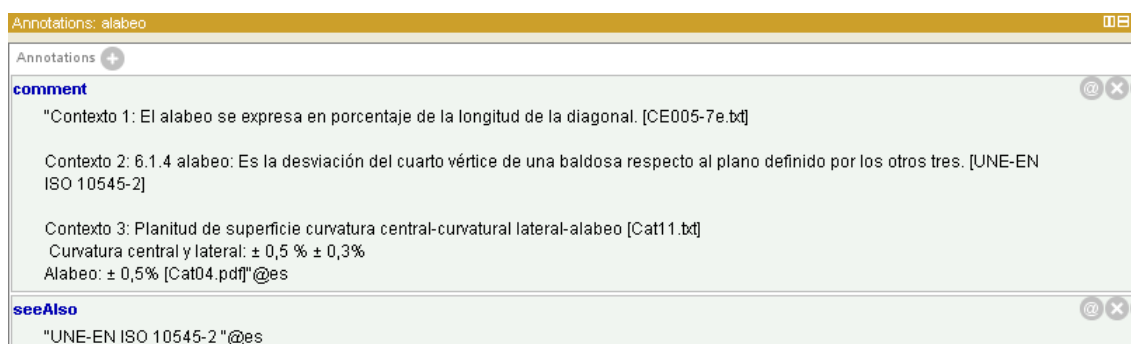


Figura 163. Contextos en la clase ‘alabeo’

Por ejemplo, en la figura superior observamos los contextos declarados en la clase ‘alabeo’ en forma de anotación de tipo ‘comment’. En cada contexto, se indica la fuente del contexto utilizando un método abreviado. Las fuentes se pueden consultar en el anexo 10.

En total, hemos introducido 131 contextos en la ontología. Todas las características aparecen en los tres tipos de recurso (corpus, normas y catálogos) a excepción de 7:

- densidad aparente
- densidad relativa aparente
- emisión de cadmio y plomo
- pequeñas diferencias de color
- resistencia al impacto
- composición
- resistencia al rayado

En el caso de ‘densidad aparente’, ‘densidad relativa aparente’, ‘emisión de cadmio y plomo’ y ‘pequeñas diferencias de color’ no hemos encontrado ocurrencias de estas características en ningún catálogo comercial de los consultados. En el caso de ‘composición’ y ‘resistencia al rayado’ no hemos encontrado ocurrencias en las normas UNE de cerámica que hemos utilizado, aunque en el caso de ‘resistencia al rayado’ sí hemos encontrado ocurrencias en la norma complementaria UNE-EN 15771 (AENOR, 2011).

7.5.4 Fuente de la característica y fuentes complementarias

La fuente de la característica se ha introducido en todas las clases de característica de la baldosa de la ontología en forma de anotación. Esta anotación se ha realizado utilizando el subelemento *seeAlso* y a continuación hemos introducido la referencia a la norma utilizando el patrón “UNE-EN/UNE-EN ISO número de norma-parte si corresponde”. Por ejemplo, en la figura 164, se muestra la referencia a la norma donde hemos encontrado la característica ‘alabeo’.

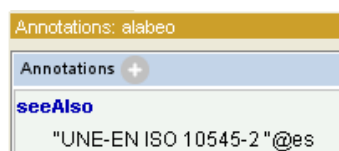


Figura 164. Anotación sobre la norma de referencia de la característica ‘alabeo’.

En el caso de las características: ‘acabado’, ‘composición’ y ‘textura’, la fuente de referencia original no ha sido una norma sino los catálogos de cerámica, en ese caso se ha indicado el código del catálogo más representativo (figura 165).



Figura 165. Anotación sobre el catálogo donde podemos encontrar la característica ‘acabado’

La introducción de esta información, junto con la de los contextos y fuentes complementarias se ha realizado de forma completa para cada característica y no por tipo de información en particular. Es decir, hemos introducido primero la información disponible en este aspecto sobre una característica y a continuación hemos pasado a la siguiente.

7.6 Descripción de las clases de baldosa

La descripción de las clases de baldosa se ha realizado en base a la información obtenida durante la fase de extracción y recogida en las fichas de baldosa (anexo 2). Como hemos mencionado anteriormente, existe una desigualdad entre el nivel de detalle de las descripciones de algunos tipos de baldosa con respecto a otras. No obstante, como la finalidad es probar si el sistema reclasifica la utilización de las características de la baldosa para describir baldosas cerámicas este aspecto no afecta en demasía. Además, en la introducción de datos hemos tenido en cuenta que, como comentamos en el apartado de formalización de la baldosa cerámica, las características distintivas entre las hermanas de una clase no siempre son las mismas para todas las hermanas de clase. Otro aspecto que hemos considerado es la descripción de las características esenciales sin limitar las clases, de manera que una baldosa que cumpla con las

características distintivas de una superclase pueda pertenecer a esta aunque no incluya en su descripción todas las características esenciales que describen a la superclase.

Durante el proceso de identificación se ha tenido en cuenta que las clases representadas como hermanas podrían pertenecer a alguna de sus hermanas de clase al agregar las descripciones de cada clase. Por esta razón, la identificación de las características distintivas de cada clase se ha basado en la comparación de las clases hermanas entre ellas y, al mismo tiempo, en las descripciones de cada clase de manera independiente, asumiendo que podría haber características que se hubieran omitido en la descripción de un tipo de baldosa pero que forman parte del tipo de baldosa.

La introducción de las descripciones de las baldosas se ha realizado por subgrupos según los hemos encontrado en los recursos textuales. En primer lugar, se han formalizado las descripciones de las clases de baldosa que representan a los grupos de la norma UNE 14411 (figura 166).

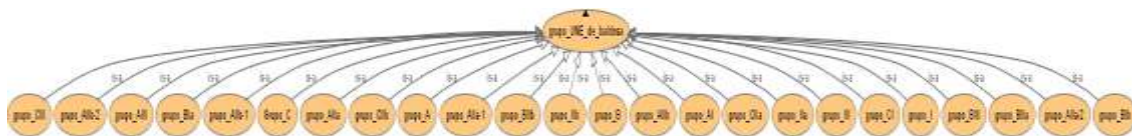


Figura 166. Clases de grupos UNE de baldosa

En el caso de las clases de grupos UNE, todas las clases de grupos se distinguen bien por la característica distintiva ‘capacidad de absorción de agua’, bien por la relación *baldosa – proceso de fabricación* o bien por la combinación de ambas excepto las que representan a los grupos AIIa-1 y AIIa-2 y los grupos AIIb-1 y AIIb-2 que comparten estos rasgos entre ellas pero se distinguen por su ‘resistencia a la flexión’.

Concretamente, las clases ‘grupo A’, ‘grupo B’ y ‘grupo C’ tienen como rasgo diferenciador la relación *baldosa – proceso de fabricación* con los valores ‘extrusión’, ‘prensado en seco’ y ‘otros procesos’, respectivamente. Las clases ‘grupo I’, ‘grupo II’ y ‘grupo III’ se distinguen por sus valores en la característica ‘capacidad de absorción de agua’. Las clases ‘grupo AI’, ‘grupo AIIa’, ‘grupo AIIb’, ‘grupo AIII’, ‘grupo BIa’, ‘grupo BIb’, ‘grupo BIIa’, ‘grupo BIIb’, ‘grupo BIII’, ‘grupo CI’, ‘grupo CIIa’, ‘grupo CIIb’ y ‘grupo CIII’ se distinguen entre ellas por sus valores en la característica ‘capacidad de absorción de agua’ y la relación *baldosa – proceso de fabricación*.

Para distinguir los grupos AIIa-1 y AIIa-2 y los grupos AIIb-1 y AIIb-2, respectivamente, la característica ‘capacidad de absorción de agua’ y la relación *baldosa – proceso de fabricación* no resultan suficientes ya que comparten valores entre ellas:

- AIIa-1 y AIIa-2 se han fabricado por extrusión y ambas tienen una capacidad de absorción de agua con valor mayor o igual 3 y menor o igual 6
- AIIb-1 y AIIb-2 se han fabricado por extrusión y ambas tienen una capacidad de absorción de agua con valor mayor que 6 y menor o igual 10

La característica que distingue estas clases entre ellas es la ‘resistencia a la flexión’ donde los valores son distintos. Además de agregar esta característica como distintiva, hemos decidido crear dos clases ‘grupo AIIa’ y ‘grupo AIIb’ para recoger esta distinción en ‘capacidad de

absorción de agua’ y ‘relación baldosa – proceso de fabricación’ que las distingue, por ejemplo, de la clase ‘grupo AI’ y la clase ‘grupo AIII’. La clase ‘grupo AIIIa’ tiene los rasgos distintivos:

- por extrusión y una capacidad de absorción de agua con valor mayor a 3 y menor o igual a 6

La clase ‘grupo AIIIb’ tiene los rasgos distintivos:

- fabricado por extrusión y una capacidad de absorción de agua con valor mayor que 6 y menor o igual a 10

Para introducir las características distintivas en cada clase, hemos editado las descripciones de la clase y hemos introducido los formalismos para característica distintiva en los apartados *Equivalent classes* y *Superclasses*. Los formalismos utilizados se describen en III.6.8.1 y coinciden con los de la descripción de características distintivas para características de la baldosa cerámica.

A continuación, mostramos tres figuras donde se observan los grandes tipos de baldosa identificados según sus características distintivas. En primer lugar, un ejemplo de clase de baldosa cuya característica distintiva es la ‘capacidad de absorción de agua’, a continuación, una clase que se distingue por sus valores en la relación *característica – proceso de moldeo* y finalmente, una clase que combina los dos rasgos para diferenciarse por la suma de estos valores del resto.



Figura 167. Clase de baldosa cerámica con característica distintiva ‘tiene capacidad de absorción de agua’

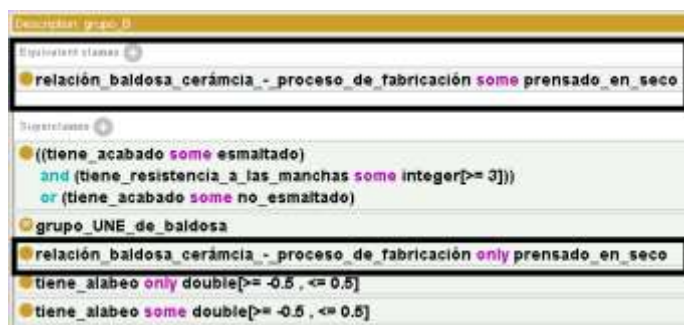


Figura 168. Clase de baldosa cerámica que se distingue por el valor en ‘relación baldosa – proceso de fabricación’



Figura 169. Clase de baldosa cerámica que se distingue por la combinación de valores en ‘tiene capacidad de absorción de agua’ y ‘relación baldosa – proceso de fabricación’

En lo que refiere a las características esenciales, se han introducido los datos según la información obtenida del análisis de grupos de tipo A y B, ya que clases ‘grupo CI’, ‘grupo CIIa’, ‘grupo CIIb’ y ‘grupo CIII’ no aparecen descritas en la norma y sólo se menciona su diferencia con respecto al resto en los valores de ‘capacidad de absorción de agua’ y ‘relación baldosa - proceso de fabricación’.

Para cada uno de los grupos hemos detectado alrededor de 15 características esenciales y muchas de estas características y sus valores son compartidos por muchas de las clases. Conscientes de que muchos de los grupos se reagruparían más tarde en clases hermanas, hemos comparado las diferencias y semejanzas entre los grupos que podría agruparse bajo las clases ‘grupo I’, ‘grupo II’ y ‘grupo III’, las diferencias y semejanzas entre los grupos que al mismo tipo podrían pertenecer al ‘grupo A’, ‘grupo B’ y ‘grupo C’ y las semejanzas entre todos los grupos descritos. Las características compartidas por todos los grupos de tipo A, B, C, I, II o III se han introducido en la superclase ‘grupos UNE de baldosa’, las características compartidas entre los grupos de tipo A se han introducido en la clase ‘grupo A’ y así, para cada grupo de los mencionados.

Por ejemplo, en la descripción de las clases ‘grupo AI’ y ‘grupo BIa’, ‘grupo BIb’ se menciona la característica ‘resistencia a la helada’ con valor positivo y otros grupos, por ejemplo el ‘grupo AIIa’ o ‘grupo AIII’ no tienen esta característica. A partir de estos datos, hemos interpretado que esta característica pertenece a todos los miembros de la clase ‘Grupo I’ y se ha introducido en esta clase.

En la introducción de las características esenciales, el complemento *Class Matrix* ha sido de gran ayuda, ya que hemos podido introducir todos los valores de una misma característica en todas las clases sin tener que editando clase a clase. A continuación se muestra un extracto de la descripción de características esenciales de estos grupos desde *Class Matrix*.

Clase	Base_datos (zona)	Forma_superficie (zona)	Espesor
(double)	(integer)	(double)	(double)
(double)	(integer)	(double)	(double)
(double)	(integer)	(double)	(double)
(double)	(integer)	(double)	(double)
double[>= -1.5 , <= 1.5] (double)	(integer)	(double)	double[>= 35] (double)
(double)	(integer)	(double)	double[>= 13] (double)
double[>= -1.5 , <= 1.5] (double)	(integer)	(double)	double[>= 9.6] (double)
(double)	(integer)	(double)	(double)
double[>= -0.8 , <= 0.8] (double)	(integer)	(double)	double[>= 20] (double)
(double)	(integer)	(double)	(double)
double[>= -1.5 , <= 1.5] (double)	(integer)	(double)	double[>= 17] (double)
(double)	(integer)	(double)	(double)
(double)	(integer)	(double)	double[>= 22] (double)
(double)	(integer)	(double)	(double)
(double)	integer[>= 85] (integer)	(double)	(double)
(double)	(integer)	(double)	(double)
double[>= -0.5 , <= 0.5] (double)	integer[>= 95] (integer)	(double)	(double)
(double)	(integer)	(double)	(double)
(double)	(integer)	(double)	(double)
(double)	(integer)	(double)	(double)
double[>= -0.8 , <= 0.8] (double)	(integer)	(double)	double[>= 23] (double)
(double)	(integer)	(double)	double[>= 30] (double)
(double)	(integer)	(double)	double[>= 35] (double)
(double)	(integer)	(double)	double[>= 18] (double)

Figura 170. Visualización de la descripción de características esenciales a través de *Class Matrix*

A continuación, se han descrito las clases de pavimento y revestimiento extraída de la clasificación de baldosas según su aplicación en la norma UNE 14411 (figura 171).

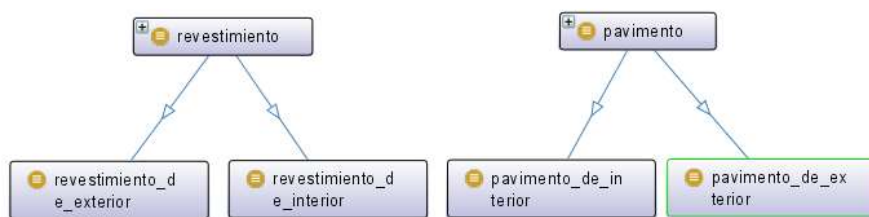


Figura 171. Clases de pavimento y revestimiento según la UNE 14411

En este caso la característica distintiva con respecto al resto de clases y entre ellas es la misma ‘tiene uso’ y lo que cambia es el valor de la característica. Estas clases tienen además características esenciales que se han introducido en la clase ‘revestimiento’ o ‘pavimento’ si también se compartían en ambas subclases ‘revestimiento de exterior’, ‘revestimiento de interior’ y ‘pavimento de interior’ y ‘pavimento de exterior’ respectivamente. Las características esenciales no compartidas por sus hermanas de clase se han formalizado en la clase concreta.

Finalmente, hemos descrito los tipos de baldosa encontrados en el manual de PROALSO (PROALSO, 2011) (figura 172).

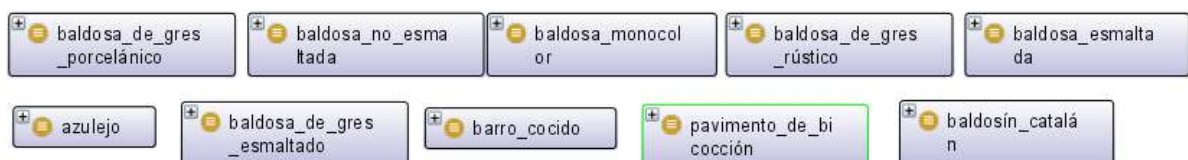


Figura 172. Clases extraídas del manual de PROALSO

La identificación de los rasgos distintivos de estas clases se ha basado en la comparación de descripciones en la ficha de baldosas y la consulta con el experto, por ejemplo, en la ‘baldosa de gres esmaltado’. La introducción de las características distintivas ha sido manual.

Para estas clases de baldosa, la introducción de características esenciales no ha considerado semejanzas o diferencias entre posibles superclases. El objetivo es que se reagrupen según lo descrito, sin asumir a qué grupo pueden pertenecer, es decir, sin identificar qué características podrían heredar de otras clases al pasar el razonador y reagrupar la clasificación.

7.7 Descripción de otros elementos en la ontología

El penúltimo paso en la fase de introducción de datos ha sido introducir las observaciones o comentarios que habíamos recogido durante la fase de extracción en los procesos, instrumentos, unidades de medida o parámetros que forman parte de la ontología. La introducción de esta información a través de anotaciones (véase III.6.5) se ha realizado por clases de concepto.

En el caso de los ‘parámetros’ se han introducido contextos, a través de la etiqueta *comment*, que ayudan a entender el significado del parámetro, la fórmula que se utiliza para calcularlo o la explicación de los valores de la característica de acuerdo con el parámetro. Hemos introducido anotaciones en cinco de los 46 parámetros de ensayo, a saber: *P*, coeficiente de absorción de agua, coeficiente de dilatación lineal, defecto visible por revoluciones y dureza de Mohs (figura 173).

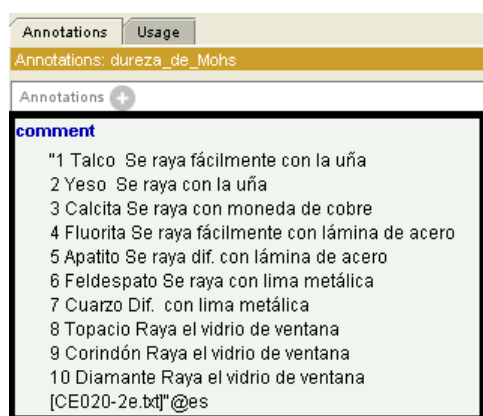


Figura 173. Comentario sobre los valores según el parámetro ‘dureza de Mohs’

En el caso de los procesos de ensayo de las características se han introducido contextos que explican cada uno de los procedimientos de limpieza (procedimiento de limpieza A, procedimiento de limpieza B, procedimiento de limpieza C, procedimiento de limpieza D). En la figura 174, mostramos la anotación a través de la etiqueta *comment* sobre el procedimiento de limpieza A.

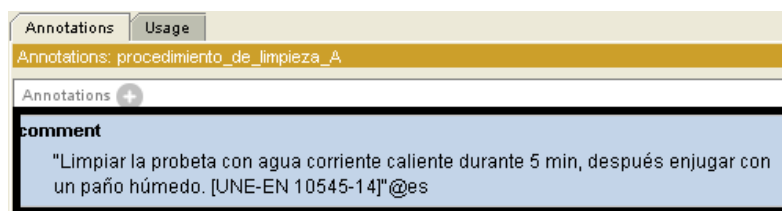


Figura 174. Comentario sobre el procedimiento de limpieza A

En el caso de los procesos de fabricación de la baldosa cerámica, hemos incluido contextos explicativos para ‘moldeo’ y ‘cocción’ a través de la etiqueta *comment*. En la figura 175 se muestra la anotación de ‘moldeo’.

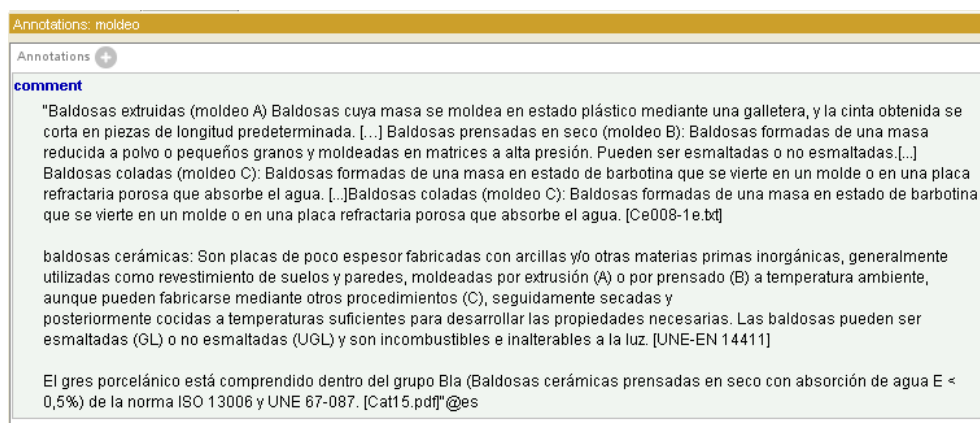


Figura 175. Comentario de la clase 'moldeo'

En las clases de 'unidad de medida' y 'instrumento de ensayo' no se han introducido anotaciones porque no hemos recogido ninguna.

7.8 Introducción de equivalentes y sinónimos

Los sinónimos y los equivalentes a otras lenguas (inglés) de las clases de características de la baldosa se han introducido desde la ventana *Annotations*, una a una, cuando teníamos la estructura de clases ya creada.

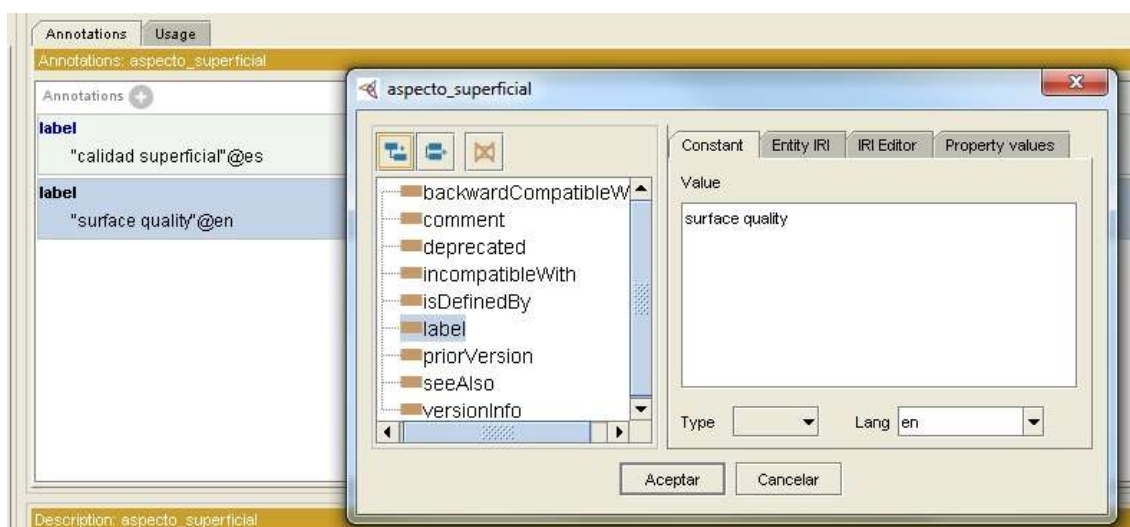


Figura 176. Introducción de equivalentes y sinónimos en forma de *annotation*.

La introducción de sinónimos y equivalentes en inglés se ha realizado en dos etapas. Primero, hemos introducido todos los equivalentes en inglés y a continuación, todas las denominaciones en español. El programa Protégé memoriza el último dato introducido en cualquiera de los apartados de la ontología, al introducir primero todos los equivalentes en inglés, hemos evitado tener que indicar cada vez el código de idioma, ya que el sistema lo tenía guardado en memoria. El procedimiento ha sido el siguiente: hemos editando cada una de las clases, hemos agregado el equivalente en el campo de anotación de la etiqueta *label* y hemos comprobado que el código de idioma era el que correspondía. Una vez introducidos todos los equivalentes en inglés, hemos repetido la operación para las denominaciones sinónimas.

7.9 Resultados de la introducción de datos en la ontología

Esta fase ha completado de manera satisfactoria la creación de la ontología de características de la baldosa cerámica. A lo largo de esta fase hemos creado 275 clases que representan a 275 conceptos en la ontología, de las cuales, 82 han sido descritas a través de alguna de las 54 propiedades de la ontología.

En las características de la baldosa (41 características + 6 clases de característica) se han introducido características esenciales y distintivas, una característica equivalente, relaciones de tipo *característica – proceso de ensayo*, relaciones de meronimia y el tipo de característica para todas las características de la baldosa cerámica (intrínseca/extrínseca).

En las baldosas cerámicas (41 tipos de baldosa) se han introducido características esenciales y distintivas y relaciones *baldosa – proceso de fabricación*.

Además, a través de anotaciones, se han introducido 131 contextos en las características de la baldosa cerámica, 5 contextos en 5 parámetros, 4 contextos en 4 procesos de ensayo de la baldosa y 3 contextos en 2 procesos de fabricación de la baldosa.

Si cuantificamos la información formalizada en base al número de elementos que conforman la ontología, el resultado, siguiendo el apartado de *Ontology Metrics* de la ontología es el siguiente:

Clases: 275	Axiomas de clase
Object property: 22	Axiomas de tipo SubClassOf: 681
Data property: 32	Axiomas de tipo EquivalentClasses: 95
Instancias: 92	Axiomas de tipo DisjointClasses: 13
Axiomas de Object Property	Axiomas de Data Property
Axiomas de tipo SubObjectPropertyOf: 19	Axiomas de tipo SubDataPropertyOf: 31
Axiomas de tipo FunctionalObjectProperty: 10	Axiomas de tipo EquivalentDataProperties: 1
Axiomas de tipo ObjectPropertyDomain: 18	Axiomas de tipo FunctionalDataProperty: 10
Axiomas de tipo ObjectPropertyRange: 18	Axiomas de tipo DataPropertyDomain: 30
Axiomas de tipo SubDataPropertyOf: 31	Axiomas de tipo DataPropertyRange: 29
Axiomas de instancia	Axiomas de Annotation Property
Axiomas de tipo ClassAssertion: 98	Axiomas de tipo AnnotationAssertion: 265
Axiomas de tipo DifferentIndividuals: 4	
Axiomas de tipo DisjointClasses: 13	

Tabla 27. Datos cuantitativos de elementos de la ontología

En conclusión, todos los datos recogidos durante la fase de extracción están presentes en la ontología.

8 La ontología de características de la baldosa cerámica

En este apartado analizamos la consistencia de la ontología y presentamos posibles consultas que se pueden realizar en la ontología para obtener información acerca de las características de la baldosa cerámica y de las baldosas.

En primer lugar, presentamos la prueba que hemos realizado en la ontología para demostrar que la formalización de las características de la baldosa cerámica es consistente. En segundo lugar, analizamos los resultados de la clasificación inferida por el razonador HermiT sobre baldosas cerámicas. Finalmente, presentamos algunos ejemplos de consultas que se pueden realizar en la ontología a través de la pestaña DL Query de Protégé.

Las consultas realizadas son una muestra de las posibilidades de búsqueda onomasiológica de la ontología.

8.1 Prueba de la inferencia y la consistencia en la formalización de características de la baldosa cerámica

En este trabajo hemos descrito formalmente las características de la baldosa cerámica y hemos estructurado las características atendiendo a las semejanzas y diferencias de estas y a su naturaleza. Las descripciones incluyen información sobre sus características esenciales y distintivas, relaciones entre características o entre características y otros conceptos, equivalencia y tipo de característica.

Para probar que las descripciones son verdaderamente consistentes, hemos realizado una sencilla operación que consiste en desvincular cada una de las características de la baldosa de sus clases y aplicar el razonador para ver si la estructura de clases inferida coincide con la estructura inicial. Explicar qué sucede si coincide y si no. Este proceso requiere editar cada una de las clases de la ontología y editar en su descripción la superclase que tiene asignada y cambiarla por ‘característica de la baldosa cerámica’. En la figura 177, se muestra un ejemplo en la descripción de la clase ‘color’ con la clase que debería eliminarse de la descripción y cambiarse por ‘característica de la baldosa cerámica’. Hemos realizado este proceso en cada una de las características. Los valores formalizados a través de clase no se han desvinculado de la característica a la que pertenecen.



Figura 177. Extracto de la descripción de 'color' y la clase que se ha eliminado de la descripción en la prueba

La estructura inicial de clases de la ‘característica de la baldosa cerámica’ se puede visualizar en la figura 178. En la figura 179, vemos cómo las características de la baldosa ya no están anidadas y se han agrupado directamente a la clase ‘característica de la baldosa’.



Figura 178. Fragmento de la estructura inicial de clases de ‘característica de la baldosa cerámica’

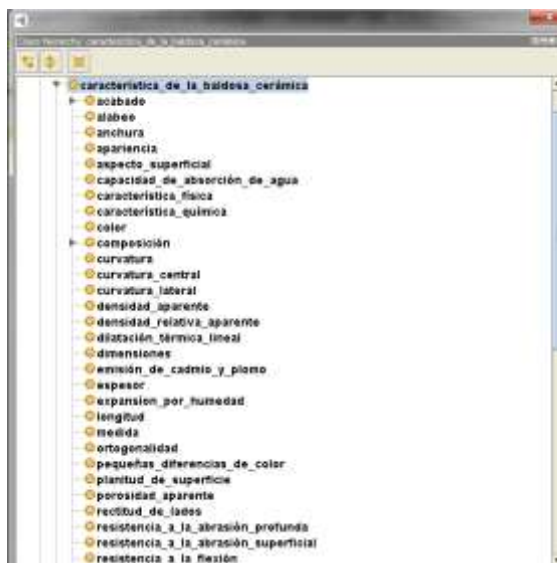


Figura 179. Extracto de la clasificación de ‘características de la baldosa cerámica’ después de desagruparlas

La aplicación del razonador HermiT a la ontología da como resultado la misma estructura inicial sólo que además las 6 subclases de ‘características de la baldosa cerámica’ se han repartido entre la clase ‘característica extrínseca’ y la clase ‘característica intrínseca’. No obstante si revisamos las características que agrupan cada una de las subclases, se demuestra que la agrupación coincide con la realizada manualmente.

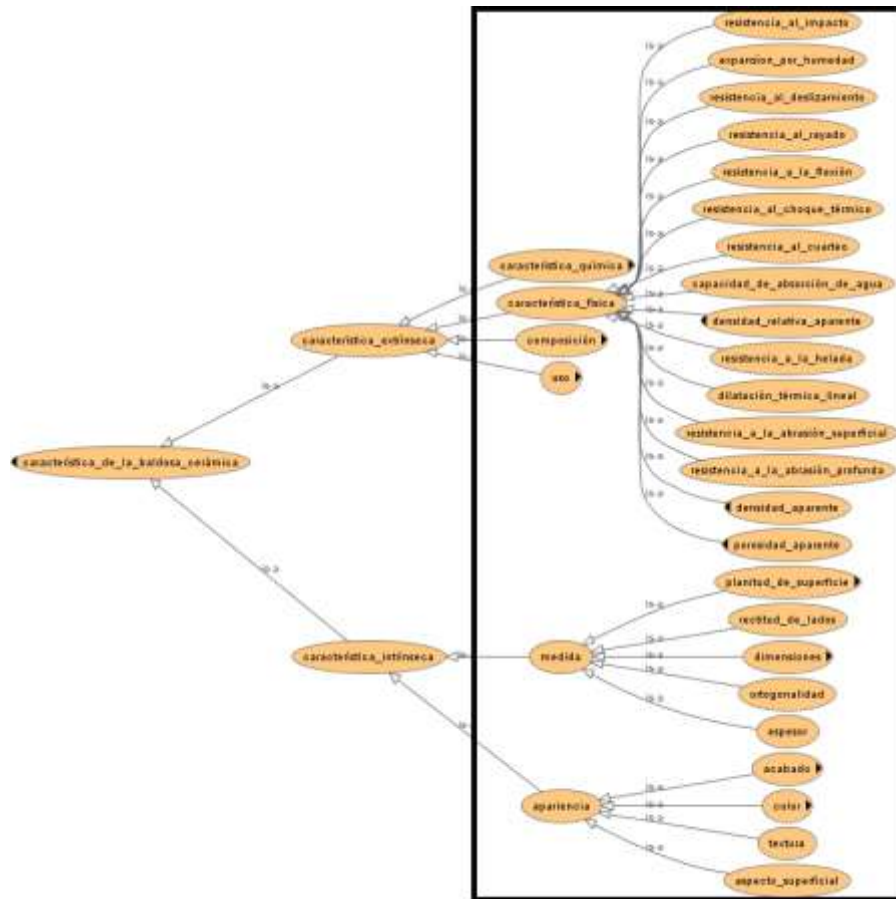


Figura 180. Grafo de las subclases inferidas de ‘característica de la baldosa cerámica’

La equivalencia entre la clasificación inicial de características de la baldosa cerámica con la clasificación inferida por el razonador HerMiT demuestra que las descripciones realizadas sobre cada característica son consistentes.

Es decir, que se ha diferenciado formalmente entre las características de la baldosa cerámica. Esto implica que la extracción de datos ha sido suficientemente exhaustiva, que la clasificación inicial es consistente y que la formalización de la información es adecuada y se corresponde con el conocimiento que queríamos representar.

8.2 Ejemplos de consultas en la ontología

La ontología describe las características de la baldosa cerámica y contiene algunas descripciones sobre tipos de baldosa cerámica. Esta ontología puede utilizarse por tanto para categorizar tipos de baldosa en sistemas informáticos adecuados. No obstante, desde la herramienta Protégé también podemos explotar algunas de las posibilidades de consulta de la ontología. Dentro de este contexto, el tipo de consultas que podemos realizar es variado y puede consistir en consultas sencillas acerca de una clase o consultas más complejas que requieren la interpretación de los resultados.

Las consultas se realizan desde la pestaña DL Query del programa y se pueden escribir utilizando la Manchester OWL 2 Syntax.

Un ejemplo de consulta sencilla puede ser: ¿qué clase de característica es ‘resistencia a las manchas’? Esta consulta se realiza escribiendo el nombre de la clase en la barra de consulta y marcando la casilla *Superclasses* en el menú derecho.

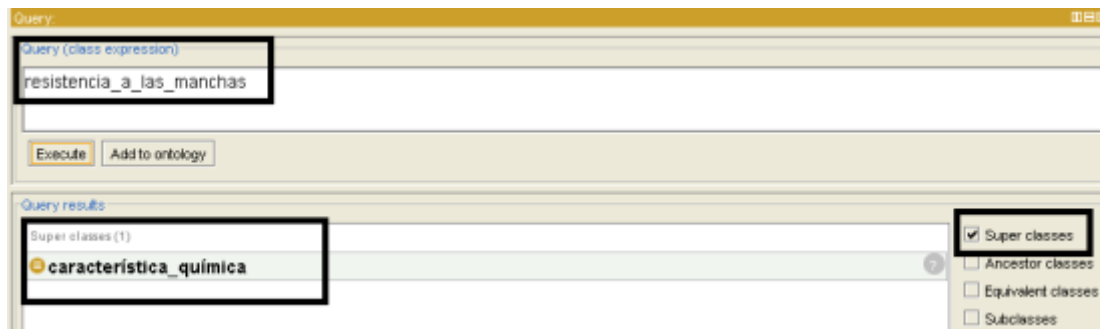


Figura 181. Resultado de la consulta sobre la clase de ‘resistencia a las manchas’

Otro tipo de consulta sencilla puede ser: ¿cuáles son los valores posibles de la característica ‘color’?. Para ello, basta con escribir el nombre de la clase y marcar la casilla de *individuals*.

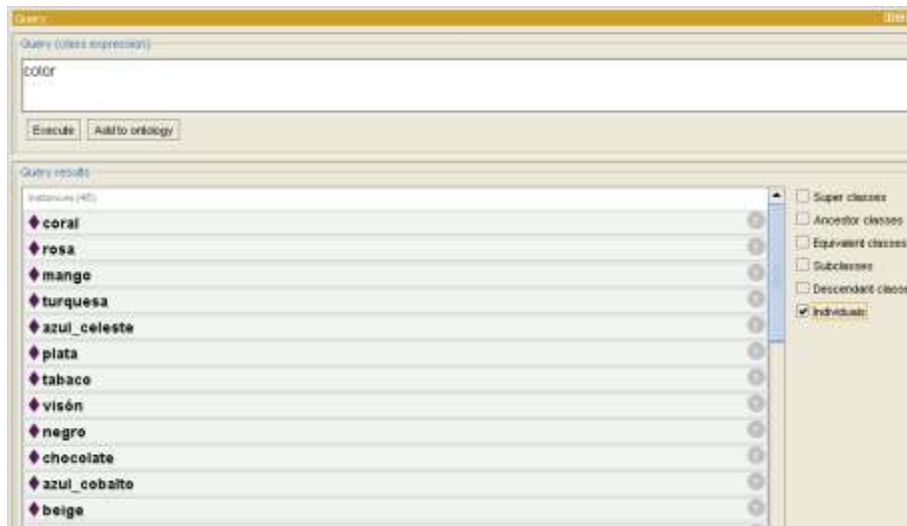


Figura 182. Extracto del resultado de la consulta sobre valores de ‘color’

Un ejemplo de consulta compleja puede ser: ¿cuál es el parámetro de ensayo de una la característica ‘capacidad de absorción de agua’? Para ello, escribimos en la barra de consulta la siguiente expresión:

`inverse tiene_parámetro_de_ensayo some capacidad_de_absorción_de_agua`

Esta expresión consulta que conceptos vinculados a capacidad de absorción de agua sirven de rango de la propiedad ‘tiene parámetro de ensayo’. En esta consulta debemos fijarnos en los resultados de la clase ‘super classes’ que incluirá las clases directamente vinculadas a esta propiedad. El resultado de la consulta se puede visualizar en la imagen a continuación y contiene los parámetros ‘coeficiente de absorción de agua’ e ‘indicadores de absorción de agua’



Figura 183. Resultado de la consulta sobre los parámetros de ensayo de 'capacidad de absorción de agua'

Finalmente, un ejemplo de consulta sobre tipos de baldosa puede ser el siguiente: ¿qué baldosas resisten a la helada? La expresión que deberíamos introducir, en este caso, sería:

`tiene_resistencia_a_la_helada value true`

Y nos fijaríamos en los resultados para *Descendant classes* (figura 184).



Figura 184. Resultado de la consulta sobre baldosas con resistencia a la helada

8.3 Análisis de la clasificación inferida de baldosas cerámicas

La formalización de tipos de baldosa cerámica en la ontología se ha realizado con dos objetivos concretos:

1. Probar la utilidad de las característica de la baldosa cerámica para la descripción de baldosas
2. Probar las posibilidades de clasificación automática de baldosas a partir de sus características

El resultado de la aplicación del razonador HermiT después de haber introducido las características de las baldosas no ha señalado ningún error de uso, lo que implica, de antemano, que la utilización de las características en la descripción de 'baldosas cerámicas' es viable.

Igualmente, no hemos notado, en la fase de introducción de datos, ninguna carencia en características o valores lo que indica que las características son suficientes para describir este tipo de ‘baldosas cerámicas’.

Para probar la descripción y clasificación de baldosas, hemos analizado los resultados de clasificación de las baldosas extraídas del manual de PROALSO. El objetivo es comprobar si de acuerdo con la descripción realizada, la baldosa se ha agrupado bajo el grupo UNE que se indicaba en el texto. Los distintos tipos de baldosa se han formalizado. El análisis del resultado después de la clasificación de baldosas mediante el razonador sugiere que las características utilizadas sirven para distinguir los conceptos pero que las descripciones requieren mejoras.

Por ejemplo, si atendíamos a la descripción de ‘baldosín catalán’ esperábamos que éste se agrupara bajo la clase ‘grupo AIIb’ y así ha sido, porque se cumplían los criterios de características para pertenecer a dicha clase (figura 185).

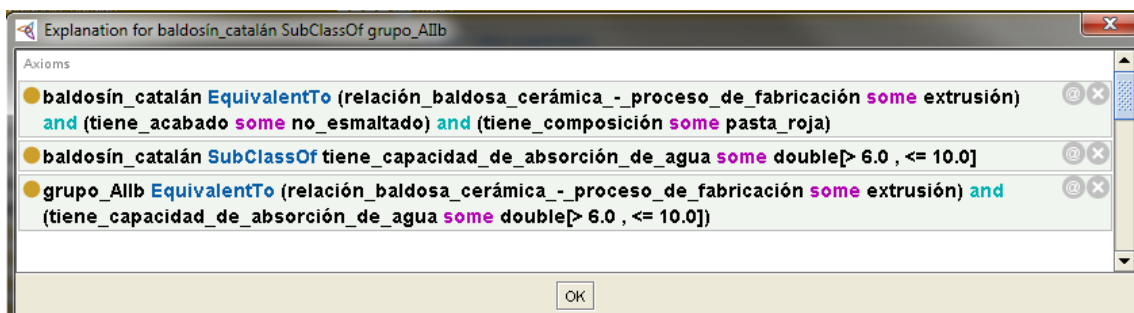


Figura 185. Explicación de ‘baldosín catalán’ como subclase de ‘grupo AIIb’

Sin embargo, en la clasificación, ‘baldosín catalán’ se ha clasificado como equivalente de ‘barro cocido’ (figura 186). Esto ha ocurrido porque la descripción de ambos tipos de baldosa coincide y sería necesario explorar con más detalle su descripción para identificar diferencias o concluir que efectivamente son equivalentes.

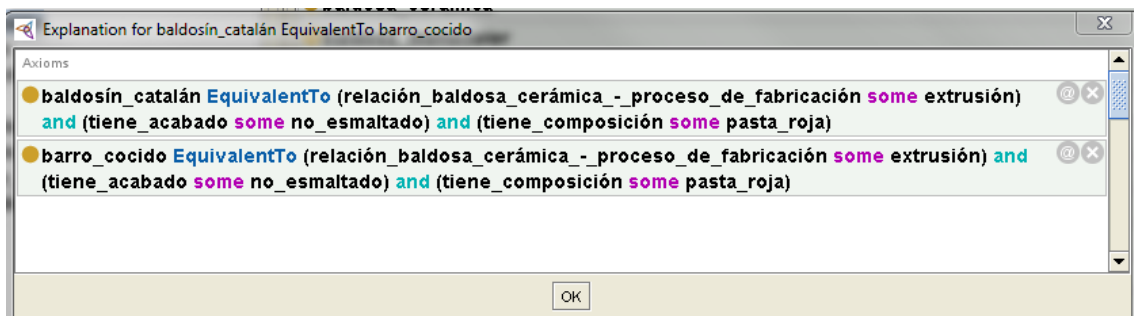


Figura 186. Explicación del razonador de la equivalencia de ‘baldosín catalán’ y ‘barro cocido’

‘Baldosa de gres rústico’ no se ha clasificado bajo el grupo AIIa, tal y como se esperaba en base a su capacidad de absorción de agua y otras características. Esto se debe a que éste exige el proceso de fabricación por extrusión como característica distintiva y en ‘baldosa de gres rústico’ no se ha declarado ninguna información acerca de esta característica. Para solucionarlo convendría agregar información sobre esta característica en la clase ‘baldosa de gres rústico’.

El resto de clases sí se han reagrupado según lo esperado.

La conclusión en este aspecto es que la utilización del razonador nos permite verificar la consistencia de una descripción. En este caso hemos comprobado que si bien las características de la baldosa formalizadas nos permiten realizar descripciones de baldosa, las descripciones realizadas para cada tipo concreto de baldosa son mejorables.

Otro aspecto interesante de aplicar el razonador a la clasificación de baldosa ha sido que hemos detectado que la clase ‘grupo BIII’, sin mencionarlo explícitamente, agrupa únicamente a baldosas esmaltadas, es decir, tiene características que sólo describen características esmaltadas. Esta clase tiene como características esenciales la ‘resistencia a la abrasión superficial’ o la ‘emisión de cadmio y plomo’ (características que describen ‘baldosas esmaltadas’, tal y como hemos descrito en sus dominios), al mismo tiempo, no tiene ninguna característica que tengan únicamente las baldosas no esmaltadas como puede ser ‘resistencia a la abrasión superficial’. Al contener únicamente características compatibles con las de las baldosas esmaltadas y no contener características propias de las baldosas no esmaltadas, el razonador infiere que “grupo BIII” agrupa únicamente a tipos de baldosa esmaltada. Habría que consultar con un experto si esta afirmación es cierta o la ontología carecía de información suficiente para hacer la clasificación de “grupo BIII” de manera adecuada.

Finalmente, en lo que respecta a los tiempos necesarios para realizar el razonamiento, el razonador tarda algo menos de un minuto y medio en realizar la clasificación de la ontología.

IV. GENERAL CONCLUSIONS

Summary and contribution

In this study, we have studied different approaches on knowledge formalization and representation and we have formalized ceramic tile properties in an ontology.

During the development of the work done to build the ontology of ceramic tile properties, we have explored a lot of ways for representing data and we have represented very different types of data. The ontology is the result of a comprehensive work that involves an amount of time and effort that it is not always available in projects oriented to commercial means. We are aware that there can be some aspects that we missed, on purpose or by mistake, but we are sure of having created an ontology whose analysis offers a large number of representation strategies, options to solve conflicts and that considers a wide range of data types.

We believe this study will contribute to the work of future developers of terminological ontologies. It will help them to determine the strategy to follow in depending on what data type to be represented, to design a formalization model according to their needs. More simply, this study contributes to the task of providing didactical contributions about how to transfer terminological knowledge to an ontology built using OWL 2 standard language.

Conclusions on the approaches and terminological considerations

Terms can be studied from different views: linguistic, cognitive or communicative. In this research and in the case of tile properties, the extraction of information has focused in texts to get to knowledge and it is been supported by the advice of an expert. In the case of tiles, we have started from the names identified in texts and tiles have been described following exclusively the criteria identified in texts.

We believe that using standards on ceramic tiles and ceramic tile properties has contributed positively to the results of this research. This type of texts is written by experts and is addressed to experts with the goal of organizing and defining the knowledge shared by the community of experts. Thus, we believe that, generally speaking, standards can be compared to the schemas that an expert would draw when describing a concept. Combining standards with other sources has allowed us to gather other views of the same field and identifying variation of terminology, it is the case of the variation of names for tiles or the differences among types of properties that had to be verified by the expert. Expert's advice has been very significant for clarifying the conceptual information, adding new knowledge to the study and to solve ambiguities.

In this study, we have tested the difference between a semasiological approach, as explained in Ontoterminology (§1.5). Here, we have used exclusively the information represented in texts in order to formalize tiles. No post-analysis of underlying knowledge neither a consulting with experts of the field has been done. We also have tested a hybrid approach, onomasiological to a large extent that combines approaching to concepts from its linguistic dimension but also from a conceptual perspective. Thus, we have not only used texts to extract terminology but have approached to the domain concepts and knowledge underlying those texts through the reading and understanding. And more important, we have been advised by an expert to validate the knowledge we were about to represent in the ontology.

Linguistic and communicative aspects have also been included in the ontology in the form of annotations.

Properties

The description of properties has considered previous studies on properties classification. Properties have been formalized according to their double role. Their formalization has been understood in way similar to the one proposed in approach such as the one of Madsen or Nuopponen: as elements denoting the qualities of other concepts and as elements being concepts themselves. Our approach pursues that properties belonging to a field of expertise should be included in terminological works considering their double role.

We have proved that concepts can be formalized according to the specific difference criteria (§1.5) and that this is not incompatible with multidimensionality. It implies describing multiple specific differences in those cases where concepts have more than one dimension in the field. For instance, in the case of ‘glazed porcelain tile’, there are two specific differences that do not occur simultaneously and depend on the coordinated concepts in the system. On the one side, it is necessary to represent the difference between ‘glazed porcelain tile’ and ‘rustic tile’ (also a type of ‘porcelain tile’). On the other side, it is necessary to distinguish between ‘glazed porcelain tile’ and ‘glazed ceramic tile’ (both being types of ‘glazed tiles’).

A definition of property suitable for our approach would be: a concept, the referent of which referent is a quality of another concept. This definition closely follows Nuopponen’s definition but opens the range of description to any kind of concept and not only entities (§2.1.4).

When distinguishing between relationships and properties we partially agree with the approach of Madsen et al (§2.1.4 y 2.1.5). From our point of view, in ontologies the representation of a relationship and a property do not differ in the form. Distinction is done in the label used to name the attribute representing the property or the relationship as considered. For instance, the domain and range of an attribute representing the relationship “process-patient” would be ‘process’ and ‘patient’ respectively. Equally, domain and range representing the property “has patient” in a process would be ‘process’ and ‘patient’ respectively. The difference relays in the focus represented in the label “process-patient” or “has patient”. We believe that the representation of a relationship among concepts and the representation of the property denoting a quality in a concept depend on the goal of the representation, on our interest for emphasizing the relationship established or the denotation of the quality. For example, in the ontology we have represented the relationship “tile-manufacturing process” emphasizing that it is a relationship. In that case our motivation was not to emphasize the property of tiles of being manufactured and describe its range but just to point that there is a relationship between tiles and the processes used to manufacture them. The possibility of completing the description of types of tile manufacturing processes and define or detail this relationship in the ontology is open to future work.

Formalization

In the regard of the methodology for building and implementing a terminological ontology, we agree with Roche et al (§1.5) that it is necessary to distinguish between the stage where the model of knowledge is designed, being this model based on a theoretical foundation of Terminology, and the stage where this model is then “moved” to a computing system using a formalization language that allows to represent it. This goal has been achieved by distinguishing between the stage where we structure, identify and limit the necessary elements that describe

tile properties and the stage where we design the model representing those elements in the ontology.

We have proposed a methodology (or strategy) to formalize in a way that concepts and terms are treated independently. The goal was emphasizing the conceptual dimension. This has been achieved by taking as a starting point the structuring of concepts; more specifically, the structuring of tile properties. In the ontology there is a structure that represents concepts (with their properties and relationships), it uses normalized terms to express them in natural language and there are also the terms referring to each concept in a language (English and Spanish) represented through the OWL 2 label 'label'. This way, the result allows us to have terms related to concepts but those terms do not interfere in the concept structure of the ontology, which remains the same regardless of the language.

Regarding ceramic tile properties, inference test results in the ontology show that data extraction has been comprehensive since the initial classification of properties proved to be consistent. Tests also show that data have been formalized correctly since there is no incoherence or ambiguity when compared to the knowledge we wanted to represent.

Furthermore, the ontology respects and accomplishes all the theoretical principles posed in the formalization model:

- It distinguishes between a property as a concept and a property when denoting a quality of another concept, and it is formalized in a way both roles are related to each other.
- Property distinction is limited according to the specific difference (§1.5) but it is modelled been flexible enough as to allow that tile properties or tiles can be classified under more than one class.
- It is language independent. It allows adding as many names of concepts as we like without having to modify the conceptual structure.
- There is a consistent formal description of tile properties. The description of properties has enough detail and it is the necessary. Enough and necessary as for ungrouping concepts of their classes and by using a reasoner having them grouped the same way on the basis of their conceptual description (shared and distinguishing properties and relationships).

Regarding the approaches uses to formalize and represent ceramic tile properties and ceramic tiles; we consider that they can contribute as examples of two types of approaches, following the distinction made by Ontoterminology. When formalizing and representing ceramic tile properties we took an approach based on non formal knowledge and addresses towards the formal representation of this knowledge. When formalizing and representing ceramic tiles the approach is based on a series of data and addresses to extract knowledge from it. This is, knowledge on ceramic tiles is not represented at the beginning but it is inferred with the help of a reasoner and the knowledge formalized through the ceramic tiles properties.

Comprehensive and shared descriptions of ceramic tiles properties and their classification as used in the formalization were considered knowledge *per se* and were represented in the ontology. In contrast, ceramic tile structure is based on data and descriptions that are exclusively based in what we explicitly found in texts. Representing these data in the ontology

pursued to obtain knowledge through the automatic reasoning of these data in combination of the knowledge already existing on ceramic tile properties in the ontology.

When comparing the initial classification of ceramic tile properties and the one inferred by the reasoner versus the initial classification of ceramic tiles and the one inferred by the reasoner we prove the significant differences existing between these representations. It also occurs when comparing the processes of formalization and the implementation stage. For instance, in the initial classification of ceramic tile properties specific difference is used to distinguish among a property and its coordinates. This suggests that conceptual space of the property and its coordinates is at the same level. After applying the reasoner in the ontology, hierarchies remain the same and no description is considered inconsistent.

In contrast, when representing ceramic tile description it was not possible to distinguish among classes by using the specific difference, which we believe is a symptom of inconsistency in the classification. This is due to the fact that in the extraction stage the classification extracted literally from the text was not exhaustive. Since we have based on what it was represented explicitly in texts, the underlying knowledge was not collected and we have tested if it was possible to infer it with the help of the reasoner.

Regarding the first approach, results are optimum. We have proved that we can formally describe ceramic tile properties and that these descriptions are consistent and comprehensive.

Regarding the second approach, results are encouraging. We have proved that using a reasoner can be very helpful to identify inconsistencies in concept classifications and descriptions; since, for instance, the terminologist could see at a glance the changes in the concept system and proceed to its validation or amendment. At the same time, we have proved that using ceramic tile properties, consistently formalized, in the description of concepts poorly described helps to increase consistency of those concepts. This is the case of the class of tile 'BIII' (§8.3). Texts did not offer any information about this class as a type of class that includes exclusively glazed tiles (and excluding unglazed ones). The use of the knowledge formalized in ceramic tile properties with domains and ranges consistently limited has led the reasoner to infer this new information.

In the respect of formalisms for representing essential and distinctive properties, we believe there is still a long way to go. In this PhD research we have explored two possibilities for formalizing essential properties. In the one, it was compulsory that members of a class have the essential property stated in their description in order to belong to that class. In the other, a concept could belong to a class if regardless of not having an essential property stated in the description, the concept did not infringe any other restriction to belong to that class; for instance, having the specific difference property which is compulsory.

The first formalism has proved to work only when description is consistent and has enough and necessary detail. This is, if concepts are not comprehensively defined or their structures are poorly limited, formalization of essential properties imposing the need of having them stated in the concept description is unfeasible. A concept lacking of this essential property in the definition (which it does not mean that the concept has not this property) would not be considered to belong to the class.

Thus, we recommend this type of formalism, used in ceramic tile properties, only if concepts of the field to be represented are comprehensively defined in advance. In contrast, the

second formalism is recommended in cases where we want to check classification consistency and infer new information. In that case, it would be more helpful to state the essential property in the class description without constraining potential members. This has been used in the description of ceramic tile classes with significant results.

Terminology and interdisciplinary approaches

This study proves the interdisciplinary harmony of Terminology and the prospects of using ontologies in Terminology. Specifically, it contributes to create terminological ontologies allowing the definition of concepts, beyond than entities, and it claims the possibilities of properties for this type of studies. Definitely, we this research we expect to contribute to increase the interest about the potentials of building and using of terminological ontologies.

Replication and future research

The ontology can be re-used to further work in the domain of ceramics. We also expect this work to contribute as a model, guide or example of representation of terminological data in ontologies. Furthermore, since the ontology is built using a standard language, it can be integrated in other knowledge-based systems.

Regarding future work, on the other hand, we believe that the strategy used to guarantee homogeneity and consistence when developing conceptual systems depends significantly on the domain. Thus, future work is related to replicate this methodology and approaches in domains where concepts refer to abstract entities (and not tangible). The goal is to test if, in these domains, schemas complexity increases and poses new needs of formalization.

On the other hand, in the basis of the results of the ontology where intrinsic properties are the most used to distinguish among concepts. We are interested in researching if this is also the case in other domains.

Related papers and contributions²⁹

Estellés, A. (2013): «Representación de terminología multilingüe en una ontología: un enfoque independiente del concepto», in *New Research in Translation and Interpreting Studies 2013*, July, Universitat Rovira i Virgili, Tarragona (Spain).

Estellés, A. (2012): «Prospects for Building an Ontology on Ceramic Tiles based on ISO Standards information», in *Toth 2012, Terminology and Ontology: Theories and Applications*, June, Annecy (France). Available at <http://www.porphyre.org/toth/actes>

Estellés, A. y A. Alcina (2012): «Ceramic Tile ISO Standards in the methodology of ontology building», in *Terminology and Knowledge Engineering Conference 2012*, June, Madrid (Spain).

Estellés, A. y A. Alcina (2009): «A model for formalizing characteristics in Protégé-OWL», in *Proceedings of the 8th International Conference on Terminology and Artificial Intelligence*, November, Toulouse (France). Available at <http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-578/>

²⁹ Online references were last checked on the 1st of December 2013

- Estellés, A. (2008): «Descripción y análisis terminológico de las características del producto cerámico acabado», in *Proceedings of the I Congreso Internacional sobre el Lenguaje de la Ciencia y la Tecnología*, October, Instituto Universitario de Lenguas Modernas Aplicadas (IULMA), Universitat Jaume I, Castellón (Spain).
- Estellés, A. (2007): «Concept characteristics in the creation of an ontology», in *Workshop Journée thématique sur la terminologie*, April, Université de Montréal, Montreal (Canada).
- Estelles, A.; A. Alcina y V. Soler (2006): «Retrieving Terminological Data from the TxtCeram Tagged Domain Corpus: First Step on a Terminological Ontology» in *Proceedings of the International Conference on Language Resources and Evaluacion*, May, Genoa (Italy). Pages 1496-1501. Available at <http://www.lrec-conf.org/proceedings/lrec2006/>
- Alcina, A. y A. Estellés (2006): «Etiquetado del corpus TxtCeram orientado a la extracción de información conceptual», in *Proceedings of XXIV Congreso internacional AESLA: Aprendizaje de lenguas, uso del lenguaje y modelación cognitiva: perspectivas aplicadas entre disciplinas*, April 2006, UNED, Madrid (Spain). Pages 1415-1424.

Bibliografía³⁰

- AENOR (1969): *UNE-ISO 1087-1: Trabajos terminológicos. Vocabulario. Parte 1: Teoría y aplicación*. Madrid, Asociación Española de Normalización y Certificación (AENOR). Modificada, vigente la edición de 2009.
- AENOR (1979): *UNE-ISO 1070-1: Trabajos terminológicos. Vocabulario. Parte 1: Teoría y aplicación*. Madrid, Asociación Española de Normalización y Certificación (AENOR). Anulada por UNE-ISO 1087-1(2009) y UNE-ISO 1087-2 (2009).
- AENOR (2000): *Clasificación Decimal Universal (CDU). Volúmenes 1 y 2: Tablas sistemáticas. Volumen 3: Índice alfabético. Edición completa*. Madrid, Asociación Española de Normalización y Certificación (AENOR).
- AENOR (2009): *UNE-ISO 1087-1: Trabajos terminológicos. Vocabulario. Parte 1: Teoría y aplicación*. Madrid, Asociación Española de Normalización y Certificación (AENOR). Vigente.
- AENOR (1991): *UNE 1066: Principios y métodos de la terminología*. Madrid, Asociación Española de Normalización y Certificación (AENOR). Vigente.
- Alcina, A. (2001): «Automatización de tareas en la elaboración de un diccionario terminológico», en *Terminología i documentació. I Jornada de Terminologia y Documentació*, Barcelona: Universitat Pompeu Fabra, 51-60.
- Alcina, A. (2006): «TXTCerám: Extracción semiautomática y análisis de términos de la cerámica a partir de un corpus electrónico. Su eficacia y utilidad en la mediación lingüística», en *XXXVI Simposio de la Sociedad Española de Lingüística*, diciembre de 2006, Madrid.
- Alonso, A; B. Guijarro; A. Lozano; J. Palma y M.J. Tabeada (2004): *Ingeniería del conocimiento. Aspectos metodológicos*. Madrid, Pearson/Prentice.
- Arntz, R. y H. Picht (1995): *Introducción a la terminología*. Madrid, Fundación Germán Sánchez Ruipérez.
- Arpírez, J.C. y Corcho, Óscar y Fernández-López, M. y Gómez-Pérez, A. (2003): «WEBODE in a nutshell». En *AI Magazine*, vol. 24, núm. 3. Págs. 37-47.
- Baader, A.F. y W. Nutt (2003): «*Basic Description Logics*». En Baader, F.; D. Calvanese; D. McGuinness; D. Nardi y P. Patel-Schneider (Eds.): *The Description Logic Handbook: Theory, Implementation and Applications*. Nueva York, Cambridge University Press. Págs. 47-100.
- Bateman, J.; B. Magnini y F. Rinaldi (1995): «The Generalized {Italian, German, French} Upper Model». En *IRST technical report*, vol. 9406, núm. 16. Trento (Italia), Istituto per la Ricerca Scientifica e Tecnologica. Págs 60-72.
- Beale, S.; S. Nirenburg y K. Mahesh (1995): «Semantic Analysis in the Mikrokosmos Machine Translation Project». En *Proceedings of the Second Symposium on Natural Language Processing (SNLP-95)*, 2-4 de agosto. Bangkok (Tailandia), Kaser Sart University. Págs. 297-307.
- Biron y Malhorta (2004): *XML Schema Part 2: Datatypes* (segunda edición). W3C [publicación electrónica]. Disponible en línea: <http://www.w3.org/TR/2004/REC-xmlschema-2-20041028/>

³⁰ Las referencias en línea se comprobaron por última vez el 1 de diciembre de 2013

- Blázquez, M.; Fernández, M.; García-Pinar, J. M., & Gómez-Pérez, A. (1998): «Building ontologies at the knowledge level using the ontology design environment». En *Proceedings of the 11th International Workshop on Knowledge Acquisition, Modeling and Management (KAW'98)*, 18-23 de abril. Alberta (Canadá), University of Calgary. Disponible en línea: http://oa.upm.es/6457/1/Building_Ontologies_at_the_K.pdf
- Bock, J.; P. Haase; Q. Ji y R. Volz (2008): «Benchmarking OWL Reasoners». En van Harmelen, F.; A. Herzig; P. Hitzler; Z. Lin; R. Piskac y G. Qi (Eds.): *Proceedings of the Workshop on Advancing Reasoning on the Web: Scalability and Commonsense*, 2 de junio, Tenerife, *CEUR Workshop Proceedings*. Disponible en línea: <http://ceur-ws.org/Vol-350/paper1.pdf>
- Bodenreider, O.; J. Willis y W. Hole (2004): «The Unified Medical Language System: What is it and how to use it?». En *MEDINFO*, 8 de septiembre de 2004, San Francisco (EEUU). Disponible en línea: http://www.nlm.nih.gov/research/umls/presentations/2004-medinfo_tut.pdf
- Borst, W. (1997): *Construction of Engineering Ontologies for Knowledge Sharing and Reuse*. Centre for Telematics and Information Technology. Enschede (Holanda), University of Twente.
- Bowker, L. (1997): «Multidimensional Classification of Concepts and Terms». En S. E. Wright y B. Gerhard (Eds.): *Handbook of Terminology Management*. Amsterdam, John Benjamins. 1: 133-143. Brachman y Schmolz, 1985)
- Brickley, D. y R. V. Guha (2004): *RDF Vocabulary Description Language 1.0: RDF Schema*. W3C [publicación electrónica]. Disponible en línea: <http://www.w3.org/TR/rdf-schema/>.
- Cabré, M. T. (1992): *La terminología. La teoría, els mètodes, les aplicacions*. Barcelona, Empúries.
- Cabré, M. T. (1993). *La terminología. Teoría, metodología, aplicaciones*. Barcelona, Antártida.
- Cabré, M. Teresa (2002). «Terminología y lingüística: la teoría de las puertas». En *Estudios de Lingüística Española*, vol.16. En línea: <http://elies.rediris.es/elies16/Cabre.html>
- Cabré, M. T.; C. Bach; R. Estopà; J. Feliu; G. Martínez y J. Vivaldi (2004): «The GENOMA-KB project: towards the integration of concepts, terms, textual corpora and entities». En *Workshop on Computational and Computer-assisted Terminology. IV International Conference on Language Resources and Evaluation (LREC 2004)*, 24-30 de mayo de 2004, Lisboa (Portugal), ELRA. Págs. 87-90. Disponible en línea: <http://www.lrec-conf.org/proceedings/lrec2004/pdf/100.pdf>
- Cadiot, P. y F. Nemo (1997a): «Pour une sémiogenèse du nom.» En *Langue Française*, vol. 113, París, Larousse. Págs. 24-34.
- Cadiot, P. y F. Nemo (1997b): « Propriétés extrinsèques en sémantique lexicale ». En *Journal of French Language Studies*, vol. 7 núm. 2, Cambridge, Cambridge University Press. Págs. 127-146.
- Campbell, K. (1990): *Abstract particulars*. Oxford (Reino Unido), Blackwell.
- Carlson, L. y S. Nirenburg (1990): *World Modeling for NPL, Technical Report* . Pittsburg (EEUU), Carnegie Mellon University.
- Clancey, W. J. (1993): «The Knowledge Level Reinterpreted: Modelling SocioTechnical Systems». En *International Journal of Intelligent Systems*, vol. 8. Págs. 33-49.
- Condamines, A. (2003): «Sémantique et Corpus Spécialisés : Constitution de bases de connaissances terminologiques». *Mémoire d'Habilitation à Diriger les Recherches*. Tolosa (Francia), Université Toulouse Le Mirail.

- Cruse, D. A. (1986): *Lexical Semantics*. Cambridge (Reino Unido), Cambridge University Press.
- Dahlberg, I. (1976): «Über Gegenstände, Begriffe, Definitionen, und Benennungen. ZurmöglichenNeufassung von DIN 2330». En *Muttersprache*, vol. 2, Mannheim (Alemania), Institut für Deutsche Sprache. Págs. 81-117.
- de Baer, P.; K. Kerremans, y R. Temmerman (2007): «A Categorisation Framework API for constructing ontology-based lexical resources». En *Proceedings of Computational Lexicography Workshop 2007*, 26-28 de octubre de 2007, Innsbrück (Austria), University of Innsbruck.
- Dentler, K.; R. Cornet; A. ten Teije y N. de Keizer (2011): «Comparison of reasoners for large ontologies in the OWL 2 EL profile». En *Semantic Web*, vol. 2, núm. 2., IOS PRESS [publicación electrónica]. Págs. 71-87. Disponible en línea: <http://iospress.metapress.com/content/n434780643m2n3u6/fulltext.pdf>
- DIN (1979): *DIN 2330: Begriffe und Benennungen; Allgemeine Grundsätze*. Berlin, Deutsches Institut für Normung (DIN). Modificada, vigente la edición de 2011.
- Drummond, N. y R. Shearer (2006): «The Open World Assumption». En *Proceedings of the Chris Date Seminar: The Closed World of Databases meets the Open World of the Semantic Web*, 12 y 12 de octubre. e-Science Institute [publicación electrónica]. Disponible en línea: <http://www.cs.man.ac.uk/~drummond/presentations/OWA.pdf>
- Drummond, N.; M. Horridge; R. Stevens; C. Wroe; S. Sampaio (2007): *Pizza Ontology v.1.15*. The University of Manchester [publicación electrónica]. Disponible en línea: <http://www.co-ode.org/ontologies/pizza/2007/02/12/>
- Estelles, A.; A. Alcina y V. Soler (2006): «Retrieving Terminological Data from the TxtCeram Tagged Domain Corpus: First Step on a Terminological Ontology». En *Proceedings of the International Conference on Language Resources and Evaluacion*, 22-28 de mayo, Génova, Italia. Págs. 1496-1501. Disponible en línea: <http://www.lrec-conf.org/proceedings/lrec2006/>.
- Estellés, A. (2012): «Prospects for Building an Ontology on Ceramic Tiles based on ISO Standards information», en *Toth 2012, Terminology and Ontology: Theories and Applications*, 7-8 de junio, Annecy (Francia). Disponible en línea próximamente en: <http://www.porphyre.org/toth/actes>
- Evans, M. W. (1988): *Relational Models of the Lexicon: Representing Knowledge in Semantic Networks*. Cambridge, Cambridge University Press.
- Faber, P. (1999): «Conceptual analysis and knowledge acquisition in scientific translation». En *Terminologie et Traduction*, vol. 2, Servicio de traducción de la Comisión Europea. Págs 97-123
- Faber, P. (2002): «ONCOTERM: sistema bilingüe de información y recursos oncológicos». En A. Alcina y S. Gamero (Eds.), *La traducción científico-técnica y la terminología en la sociedad de la información*. Universitat Jaume I, Castellón. Págs. 177-188.
- Faber, P. (2009): «The cognitive shift in terminology and specialized translation». En *MonTI. Monografías de Traducción e Interpretación*, núm. 1, Universidad de Alicante, Universidad Jaume I de Castellón y Universidad de Valencia. Págs. 107-134. Disponible en línea: <http://rua.ua.es/dspace/handle/10045/13039>
- Faber, P. y P. León (2010): «Dinamismo conceptual en las bases de conocimiento terminológico: el caso de EcoLexicon». En *Íkala, revista de lenguaje y cultura*, vol. 15, núm. 25, Universidad de Antioquia. Págs 75-100.
- Felber, H. y H. Picht (1984): *Métodos de terminografía y principios de investigación terminológica*. Madrid, Instituto "Miguel de Cervantes", CSIC.

- Feliu, J. (2000): *Relacions conceptuals i variació funcional: Elements per a un sistema de detecció automàtica*. Barcelona, Universitat Pompeu Fabra (tesina doctoral).
- Feliu, J. y M. T. Cabré (2002): «Conceptual relations in specialized texts: new tipology and an extraction sistema proposal». En *Proceedings of the 6th International Conference: Terminology and Knowledge Engineering*, 28- 30 de agosto, Nancy (Francia), INRIA. Págs. 45 -49. Disponible en línea: <http://www.upf.edu/pdi/df/teresa.cabre/docums/ca02fe.pdf>.
- Feliu, J. (2004): *Relacions conceptuals i terminologia: anàlisis i proposta de detecció automàtica*. Barcelona, Universitat Pompeu Fabra (tesis doctoral). Disponible en línea: <http://www.iula.upf.edu/publi038.htm>
- Feliu, J.; J. J. Giraldo; V. Vidal y M.T. Cabré (2004): «The Genoma-KB Project: A Concept Based Term Enlargement System». En *Proceedings of the Workshop on Computational and Computer-assisted Terminology, International Conference on Language Resources and Evaluacion*, 24-30 de mayo, Lisboa, Portugal. Págs. 32-35. Disponible en línea: <http://www.upf.edu/pdi/df/teresa.cabre/docums/ca04fel.pdf>
- Fellbaum, C. (1998): *WordNet: An Electronic Lexical Database*. Massachusetts (EEUU), MIT Press.
- Fillmore, C. J. & S. Atkins (1992): «Towards a frame-based organization of the lexicon: The semantics of RISK and its neighbors». En A. Lehrer y E. Kittay (Eds.): *Frames, Fields, and Contrast: New Essays in Semantics and Lexical Organization*, Hillsdale (EEUU), Lawrence Erlbaum. Págs. 75-102.
- Fillmore, C. J. (1982): «Frames and the Semantics of Understanding». En *Quaderni di Semántica*, vol. 6, núm. 2, Bolonia (Italia), Società editrice il Mulino. Págs.224-254. Versión disponible en línea: <http://www.icsi.berkeley.edu/pubs/ai/framesand85.pdf>
- Gaines, B.R. (1990): «Organizational Knowledge». En B. Gaines y J. Hoose (Eds.), *The foundations of knowledge acquisition*, Calgary (Canadá), University of Calgary. Págs. 192-203
- Gärdenfors, P. (2004): *Conceptual spaces: the geometry of thought*. Massachusetts (EEUU), MIT Press.
- Gaudin, F. (2000): «Les termes ont-ils des propriétés extrinsèques?». En H. Béjoint y P. Thoiron (Eds.), *Le sens en Terminologie*, Lyon (Francia), Presses universitaires de Lyon. Págs. 153-181.
- Geentjens, S.; K. Kerremans; P. de Baer y R. Temmerman (2006): «Sociocognitive terminology and terminography». En *Journées d'Etudes sur le Traitement Automatique de la Langue Arabe*, 5 y 6 de junio de 2006, Rabat (Marruecos), Institut d'Etudes et de Recherches pour l'Arabisation. Págs. 138-151.
- Goldberg, A. (1998): «Patterns of experience in patterns of language». En Tomasello (Ed.), *The New Psychology of Language. Cognitive and Functional Approaches to Language Structure*. Nueva Jersey/Londres, Erlbaum. 203-220.
- Gómez- Pérez, A.; M. Fernández-López y O. Corcho (2004): *Ontological Engineering*. Londres (Reino Unido), Springer Verlag.
- Goodman, N. (1951): *The Structure of Appearance*. Cambridge (EEUU), Harvard University Press.
- Gruber, T. (1993): «A translation approach to portable ontology specification». En *Knowledge Acquisition*, vol. 5, núm. 2, Stanford (EEUU), Stanford University Press. Págs. 199-220. Disponible en línea: <http://tomgruber.org/writing/ontolingua-kaj-1993.htm>
- Gruber, T. (1995): *Toward Principles for the Design of Ontologies Used for Knowledge Sharing*. En *International Journal Human-Computer Studies*, vol. 43, núm. 5-6, Amsterdam (Holanda), Elsevier Science Publishers. Págs. 907-928. Disponible en línea: <http://tomgruber.org/writing/onto-design.htm>

- Guarino, N. (1995): «Formal Ontology, Conceptual Analysis, and Knowledge Representation». En *International Journal of Human and Computer Studies*, vol. 43, núm. 5-6, Amsterdam (Holanda), Elsevier Science Publishers. Págs. 625-640. Disponible en línea: <http://www.loa.istc.cnr.it/Papers/FormOntKR.pdf>
- Guarino, N. y P. Giaretta (1995): «Ontologies and Knowledge Bases: Towards a Terminological Clarification». En *Towards Very Large Knowledge Bases: Knowledge Building and Knowledge Sharing*. Amsterdam (Holanda), IOS PRESS. Págs. 25 – 32.
- Hayes P.J (1979): «The logic of frames». En D. Metzger (ed.), *Frame Conceptions and Text Understanding*. Berlín, Walter de Gruyter and Co. Págs. 46–61.
- Hitzler, P.; M. Krötzsch; B. Parsia; P. Patel-Schneider; S. Rudolph (2009): *OWL 2 Web Ontology Language. Primer*. W3C [publicación electrónica]. Disponible en línea: <http://www.w3.org/TR/2009/REC-owl2-primer-20091027/>
- Horridge (2011) A Practical Guide To Building OWL Ontologies Using Protégé 4 and CO-ODE Tools. Edition 1.3. The University of Manchester [publicación electrónica]. Disponible en línea: http://owl.cs.manchester.ac.uk/tutorials/protegeowltutorial/resources/ProtegeOWLTutorialP4_v1_3.pdf
- Horridge, M. y P. Patel-Schneider (2009): *OWL 2 Web Ontology Language Manchester Syntax*, W3C [publicación electrónica]. Disponible en línea: <http://www.w3.org/TR/2009/NOTE-owl2-manchester-syntax-20091027/>
- Horridge, M.; N. Drummond; J. Goodwin; A. L. Rector; R. Stevens; H. Wang (2006): «The Manchester OWL Syntax». En B. Cuenca, P. Hitzler, C. Shankey y E. Wallace (Eds.): *Proceedings of the OWLED '06 Workshop on OWL: Experiences and Directions*, Athens (EEUU), 10-11 de noviembre de 2006, CEUR-WS.org [publicación electrónica]. Disponible en línea: http://ceur-ws.org/Vol-216/submission_9.pdf
- ISO (1995): *ISO 1087-1 Terminology work - Vocabulary. Part 1: Theory and application*. Ginebra, International Organization for Standardization (ISO). Modificada en 2000.
- ISO (1999): *ISO 12620: Computer applications in terminology. Data categories*. Ginebra, International Organization for Standardization (ISO). Modificada en 2009.
- ISO (2000): *704: Terminology work. Principles and methods*. Ginebra, International Organization for Standardization (ISO). Modificada en 2009.
- Jiménez-Ruiz, E (2010): *Logic-based support for Ontology Development in Open Environments (Tesis doctoral)*. Universitat Jaume I, Departament de Llenguatges i Sistemes Informàtics. Disponible en línea: <http://krono.act.uji.es/PhDs/DissertationErnesto.pdf/view>
- Jimeno-Yepes, A., E. Jiménez-Ruiz, R. Berlanga-Llavori y D. Rebolz-Schuhmann (2009): «Reuse of terminological resources for efficient ontological engineering in Life Sciences». En *BMC Bioinformatics 2009*, vol 10, num 4. En línea: <http://www.biomedcentral.com/content/pdf/1471-2105-10-S10-S4.pdf>
- Kageura, K. (1997): «Multifaceted/Multidimensional Concept Systems». En S. E. Wright y G. Budin (Eds.): *Handbook of Terminology Management*, vol. 1. Amsterdam (Holanda), John Benjamins. Págs. 119-132.
- Kageura, K. (2002): *The Dynamics of Terminology: a descriptive theory of term formation and terminological growth*. Amsterdam (Holanda), John Benjamins.
- Kalyanpur, A.; B. Parsia; E. Sirin; B. Grau y J. Hendler (2006): «Swoop: A Web Ontology Editing Browser». En *Journal Web Semantics: Science, Services and Agents on the World Wide Web archive*, vol. 4, núm. 2. Amsterdam (Holanda), Elsevier Science Publishers. Págs. 144-153.

- Kerremans, K.; P. de Baer y R. Temmerman (2007): «Dealing with terminological variation in Termonography: examples from the PoCeHRMOM project». En *XVIth European Symposium on Language for Special Purposes*, 27-31 de agosto de 2007. Hamburgo (Alemania), Peter Lang. Disponible en línea: http://taalkunde.ehb.be/sites/www2.ehb.be/files/u96/LSP07_KPR.pdf
- Kerremans, K.; R. Temmerman y P. de Baer (2008): «Constructing domain knowledge via terminological understanding». En *Linguistica Antverpiensia*, vol. 7. Artesis University College Antwerp [publicación electrónica]. Págs. 177-191. Disponible en línea: <http://www.lans-tts.be/img/NS7/Kerremans.pdf>
- Kocourek, R. (1991): *La langue française de la technique et de la science. Vers unelinguistique de la languesavante*. Wiesbaden (Alemania), BrandstetterVerlag (2ª ed.).
- Langacker, Ronald (1991): *Concept, Image and Symbol. The Cognitive Basis of Grammar*. Berlín, Mouton.
- Lakoff, G. (1987): *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago, University of Chicago.
- Lassila, O. y D. McGuinness (2001): «The Role of Frame-based Representation on the Semantic Web». En *Linköping Electronic Articles in Computer and Information Science*, vol. 6. Núm. 5. Linköping (Suecia), LiU Electronic Press. Disponible en línea: <http://www.ep.liu.se/ea/cis/2001/005/>.
- L'Homme, M.C.; U. Heid y J. C. Sager (2003): «Terminology during the past decade (1994–2004)». En *Terminology*, vol. 9, núm. 2. Amsterdam (Hollanda), John Benjamins. Págs. 151–161
- López, C. I.; M. Tercedor y P. Faber (2006): «Gestión terminológica basada en el conocimiento y generación de recursos de información sobre el cáncer: el proyecto Oncoterm». En *Revista eSalud*, vol. 2, núm. 8, Fundación para la eSalud [publicación electrónica]. Disponible en línea: <http://www.revistaesalud.com/index.php/revistaesalud/article/view/127/325>
- Luther, M.; T. Liebig; S. Boehm y O. Noppens. «Who the Heck is the Father of Bob?». En *ESWC 2009 Heraklion Proceedings of the 6th European Semantic Web Conference on The Semantic Web: Research and Applications*, Berlin, Heidelberg (Alemania), Springer-Verlag. Págs. 66 – 80.
- Lyons, J. (1977): *Semantics*. Cambridge (Reino Unido), Cambridge University Press.
- Madsen, B. N. (1998): Typed featured structures for terminology work - Part I. LSP - Identity and Interface - Research, Knowledge and Society. En L. Lundquist, H. Picht y J. Quistgaard (Eds.), *LSP, identity and interface. Research, knowledge and society. Proceedings from the 11th European Symposium on LSP, 1997*, vol. 1. Copenague (Dinamarca), Copenhagen Business School. Págs. 339-348
- Madsen, B. N. y A. E. Odgaard (2010): «From Concept Models to Conceptual Data Models». En *Proceedings of the Terminology and Knowledge Engineering Conference (TKE 2010)*, 12 -13 de agosto. Dublín (Irlanda), Dublin City University. Págs. 537- 544. Disponible en línea: <http://openarchive.cbs.dk/bitstream/handle/10398/8284/bnm-aeo-TKE-2010-NEW.pdf?sequence=1>
- Madsen, B. N., H. E. Thomsen y C. Vinker (2005): «Multidimensionality in Terminological Concept Modelling». En B. N. Madsen y H. E. Thomsen (Eds.): *Terminology and Content Development*, Copenague (Dinamarca), Litera. Págs. 161-174.
- Madsen, B. N.; H. E. Thomsen y C. Vikner (2002): «The CAOS Project - Computer Aided Ontology Structuring. Foundations and Applications of Conceptual Structures» . En G. Angelova, D. Corbett y U. Priss(Eds.): *Contributions to ICCS 2002*, Borovets (Bulgaria), Bulgarian Academy of Sciences. Págs. 29-32.
- Manola, F. y E. Miller (2004): *RDF Primer*. W3C [publicación electrónica]. Disponible en línea: <http://www.w3.org/TR/2004/REC-rdf-primer-20040210/>.

- Maroto, N. (2007): *Las relaciones conceptuales en la terminología de los productos cerámicos y su formalización mediante un editor de ontologías*. Castellón, Universitat Jaume I. (Tesis doctoral).
- Masolo, C.; S. Borgo; A. Gangemi; N. Guarino y A. Oltramari (2003): *WonderWeb EU Project Deliverable D18*. Trento (Italia), ISTC-CNR [publicación electrónica]. Disponible en línea: <http://wonderweb.semanticweb.org/deliverables/documents/D18.pdf>
- McGuinness, D. y F. Van Harmelen (2004): *OWL Ontology Web Language Overview*. W3C [publicación electrónica]. Disponible en línea: <http://www.w3.org/TR/2004/REC-owl-features-20040210/>
- Meyer, I. y K. Macintosh (2000): «When terms move into our everyday lives: An overview of de-terminologization». En *Terminology*, vol. 6, núm. 1, Amsterdam (Holanda), John Benjamins. Págs. 111-138.
- Meyer, I.; D. Skuce; L. Bowker y K. Eck (1992a): «Towards A New Generation Of Terminological Resources: An Experiment In Building A Terminological Knowledge Base». En *Proceedings of COLING 1992, 14th International Conference on Computational Linguistics*, 23-28 de agosto, Nantes (Francia). Págs. 956-960. Disponible en línea: <http://acl.ldc.upenn.edu/C/C92/C92-3146.pdf>
- Meyer, I.; E. Karen y D. Skuce (1997): «Systematic Concept Analysis within a Knowledge-Based Approach to Terminology». En S. E. Wright y G. Budin (Eds.): *Handbook of Terminology Management*, vol. 1. Amsterdam, John Benjamins. Págs 98-118.
- Meyer, I.; L. Bowker y K. Eck (1992b): «COGNITERM: An Experiment in Building a Terminological Knowledge Base». En H. Tommola, K. Varantola, T. Salmi-Tolonen y J. Schopp (Eds.): *Euralex '92, Proceedings I-II. Papers Submitted to the 5th Euralex International Congress on Lexicography in Tampere, Finland, vol 1*. Págs. 161- 172. Disponible en línea: http://www.euralex.org/elx_proceedings/Euralex1992_1/
- Meyer, I. 1993. «Concept Management for Terminology: a Knowledge Management Approach». En R. A. Strehlow y S. E. Wright: *Standardizing Terminology for Better Communication: Practice, Applied Theory and Results*, Filadelfia (EEUU), American Society for Testing and Materials. Págs. 140-151.
- Mignorance, M. L. (1995): «Lexical Logic and Structural Semantics: Methodological Underpinnings in the Structuring of a Lexical Database for Natural Language Processing». En U. Hoinkes (Ed.): *Panorama der Lexikalischen Semantik*. Tübingen, Gunter Narr. Págs. 461-474.
- Mignorance, M. L. (1984): «Lexical Fields and Stepwise Lexical Decomposition in a Contrastive English-Spanish Verb Valency Dictionary». En R. Hartmann (Ed.): *LEXeter 83. Proceedings of the International Conference on Lexicography*. Tübingen, Niemeyer. Págs. 226-236.
- Miller, G. (1998): «Nouns in Wordnet». En Fellbaum, C. (Ed.): *WordNet: An Electronic Lexical Database*. Massachusetts (EEUU), MIT Press. Págs. 23-46.
- Miller, G. A.y C. Fellbaum (1991): «Semantic Networks of English». Reedición en B. Levin y S. Pinker (Eds.): *Lexical and Conceptual Semantics*, Cambridge (EEUU), Blackwell. Págs. 197-229.
- Minsky (1975): «A Framework for Representing Knowledge». En P.H. Winston (Ed.): *The Psychology of Computer Vision*. Nueva York, McGraw-Hill.
- Mizoguchi, R.; J. Vanwelkenhuysen y M. Ikeda (1995): «Task Ontology for Reuse of Problem Solving Knowledge». En N. Mars (Ed.), *Towards Very Large Knowledge Bases*, Amsterdam (Holanda), IOS Press. Págs. 46-59. Disponible en línea: <http://www.ei.sanken.osaka-u.ac.jp/pub/miz/KBKS95.pdf>
- Monterde, A. M. (2002): «Relationship and dependency between linguistic and non-linguistic forms of concept representation: A study of texts addressed to experts and students». En *LSP & Professional*

- Communication*, vol. 2, núm.2. Copenague (Holanda), Dansk Selskab for Fagsprog of Fagkommunikation. Págs. 31-48.
- Moreno, A. J. (2002): «OntoTerm: un sistema abierto de representación conceptual». En *Actas del XVI Congreso de la Sociedad para el Procesamiento del Lenguaje Natural*, 26-28 de septiembre de 2002. Vigo, SEPLN.
- Motik, B.; P. F. Patel-Scheider y B. Parsia (Eds.) (2009): *OWL 2 Web Ontology Language: Structural Specification and Functional-Style Syntax*. W3C, [publicación electrónica]. Disponible en línea: <http://www.w3.org/TR/2009/REC-owl2-syntax-20091027/>
- Nilsson, J. F. (2005): «Approaches to modelling of classes and properties». En *Workshop on Modelling Classes & Properties in Terminology & Ontologies, 7th International Conference on Terminology and Knowledge Engineering*, 16 de agosto de 2005. Copenague (Dinamarca), Copenhagen Business School.
- Noy, N. F. y D. L. McGuinness (2005): *Ontology Development 101: A Guide to Creating Your First Ontology*, Stanford Knowledge Systems Laboratory Technical Report KSL-01-05. Stanford, University of Stanford [publicación en línea]. Disponible en línea: <http://www.ksl.stanford.edu/people/dlm/papers/ontology-tutorial-noy-mcguinness.pdf>
- Noy, N. F., R. Ferguson, y M. A. Musen (2000): «The knowledge model of Protégé-2000: Combining interoperability and flexibility». En *Proceedings of 12th International Conference, EKAW 2000. Knowledge Engineering and Knowledge Management Methods, Models, and Tools*, 2-6 de octubre de 2000, Juan-les-Pins (Francia). Berlin, Springer-Verlag. Págs. 17-32.
- Nuopponen, A. (1994): «On causality and concept relationships». En J. K. Draskau y H. Picht (Eds.): *Terminology Science and Terminology Planning, IITF-Workshop on Theoretical Issues of Terminology Science*. Vienna, TermNet. Págs. 217-230.
- Oltramari, A.; A. Gangemi; N. Guarino y C. Masolo (2002): «Restructuring WordNet's Top-Level: The OntoClean approach». En K. Simov (Ed.): *Workshop Proceedings of OntoLex'2, Ontologies and Lexical Knowledge Bases*, 27 de mayo, Las Palmas, España. Sofia, OntoText Lab. Págs. 17-26. Disponible en línea: <http://www.bultreebank.org/OntoLex02/OntoLex02Paper03.pdf>
- Otman, G. (1996): *Les représentations sémantiques en terminologie*. París (Francia), Masson.
- Patel-Schneider, P.; P. Hayes, y I. Horrocks (2004): *OWL Web Ontology Language Semantics and Abstract Syntax*. W3C [publicación electrónica]. Disponible en línea: <http://www.w3.org/TR/2004/REC-owl-semantics-20040210/>
- Picht, H. y J. K. Draskau (1985): *Terminology: an introduction*. Guilford (Reino Unido), University of Surrey.
- Prieto, J. A y C. I. López (2009): «Managing graphic information in terminological knowledge bases». *Terminology*, vol. 15, num. 2. Págs. 179-213.
- Pugh, J. M. (1984): *A Contrastive Conceptual Analysis and Classification of Complex Noun Terms in English, French and Spanish with Special Reference to the Field of Data Processing*. Manchester (Reino Unido), University of Manchester. (Tesis doctoral)
- Quillian, M. R. (1967): «Words concepts: A theory and simulation of some basic semantic capabilities». En *Behavioral Science*, vol. 12, núm. 5. Págs. 410-430.
- Rich y Knight (1991): *Artificial Intelligence*. New York, McGraw-Hill.
- Roche C. (2007): «Saying is not modelling». En *Proceedings of NLPCS 2007, Natural Language Processing and Cognitive Science*, Funchal, Portugal, 12-13 de junio. Págs. 47 – 56. Disponible en línea: <http://ontology.univ-savoie.fr/condillac/files/docs/articles/2007-nlpcs-saying-is-not-modelling.pdf>

- Roche, C. (2001): «The ‘Specific-Difference’ Principle: a Methodology for Building Consensual and Coherent Ontologies». En *Proceedings of the International Conference on Artificial Intelligence (IC-AI 2001)*, 25-28 de junio, Las Vegas (EEUU).
- Roche, C. (2005): «Terminologie et ontologie». En *Langages*, vol. 39, núm. 157. Págs. 48-62. Disponible en línea: http://www.persee.fr/web/revues/home/prescript/article/Igge_0458-726x_2005_num_39_157_974
- Roche, C. (2008): «Le terme et le concept: fondements d'une ontoterminologie». En *Terminologie et Ontologie: Théories et Applications (TOTH 2007)*, 1 y 2 de junio, Annecy (Francia), InstitutPorphyre. Disponible en línea: <http://arxiv.org/ftp/arxiv/papers/0801/0801.1275.pdf>
- Roche, C.; M. Calberg-Challot; L. Damas, P. Rouard (2009): «Ontoterminology: A new paradigm for terminology». En *KEOD 2009. International Conference on Knowledge Engineering and ontology Development*, 5-8 de octubre de 2009, Madeira (Portugal). Disponible en línea: <http://ontology.univ-savoie.fr/condillac/files/docs/articles/Ontoterminology-a-new-paradigm-for-terminology.pdf>
- Rosch, E. (1978): «Principles of categorization». En E. Rosch y B. Lloyd (Eds.), *Cognition and categorization*. Nueva Jersey, Erlbaum. Págs. 27-48.
- Sager, J. C. (1993): *Curso práctico sobre el procesamiento de la terminología*. Madrid, Fundación Germán Sánchez Ruipérez.
- Shearer, R.; M. Boris y I. Horrocks. *HermiT: A Highly-Efficient OWL Reasoner*. En A. Ruttenberg, U. Sattler y C. Dolbear (Eds.): *Proceedings of the 5th International Workshop on OWL: Experiences and Directions (OWLED 2008 EU)*, 26-27 de octubre de 2008, Karlsruhe (Alemania). Disponible en línea: <https://www.cs.ox.ac.uk/people/boris.motik/pubs/smh08HermiT.pdf>
- Skuce, D. y T. Lethbridge (1995): «CODE4: a unified system for managing conceptual knowledge». En *International Journal of Human and Computer Studies*, vol. 42, núm. 4. Amsterdam (Holanda), Elsevier Science Publishers. Págs. 413-451
- Soler, V., A. Alcina y A. Estellés (2006): «La digitalización de textos para la elaboración de un corpus lingüístico electrónico: una experiencia de trabajo en equipo con estudiantes». En *X Jornades de Traducció i Interpretació a Vic: Tecnologies a l'abast*, 30-31 de marzo de 2006, Universitat de Vic. ISBN: 84-934995-2-8.
- Sowa, J. F. (1999): *Knowledge Representation: Logical, Philosophical and Computational Foundations*. Pacific Grove (EEUU), Brooks Cole Publishing Co.
- Sure, Y.; M. Erdmann; J. Angele; S. Staab y R. Studer; D. (2002): «Ontoedit: Collaborative Ontology Engineering for the Semantic Web». En *First International Semantic Web Conference 2002*, Cerdeña (Italia), Springer-Verlag. Disponible en línea: <http://userpages.uni-koblenz.de/~staab/Research/Publications/iswc2002sub.pdf>
- Sure, Y.; M. Erdmann; J. Angele; S. Staab; R. Studer y D. Wenke (2002): «Ontoedit: Collaborative Ontology Engineering for the Semantic Web». En *Proceedings of the First International Semantic Web Conference*, Berlin, Springer-Verlag. Págs. 221-235. Disponible en línea: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.111.7862>
- Studer R.; R. Benjamins y D. Fensel (1998): «Knowledge engineering: Principles and methods». En *Data & Knowledge Engineering*, vol. 25, núm 1-2. Amsterdam (Holanda), Elsevier Science Publishers. Págs. 161-198.
- Swartout, B. y Y. Gil (1995): «EXPECT: Explicit Representations for Flexible Acquisition». En *Proceedings of the Ninth Knowledge Acquisition for Knowledge-Based Systems (KAW '95)*, 10-13 de abril, Alberta

(Canadá), *University of Calgary*. Disponible en línea: <http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA308790>

Temmerman, R. (2000): *Towards New Ways of Terminology Description: the sociocognitive approach*. Amsterdam, John Benjamins.

Temmerman, R. y K., Koen (2003): Termontography: Ontology Building and the Sociocognitive Approach to Terminology Description. En *Proceedings of CIL17*, Praga, Matfyzpress. Disponible en línea: http://taalkunde.ehb.be/sites/www2.ehb.be/files/u96/temmerman_art_prague03.pdf

TERMCAT (2009): *La definició terminològica*. Vic, Eumo Editorial.

Thomsen, H. E. (1998): «Typed featured structures for terminology work - Part II. LSP - Identity and Interface - Research, Knowledge and Society». En L. Lundquist, H. Picht y J. Quistgaard (Eds.): *LSP, identity and interface. Research, knowledge and society. Proceedings from the 11th European Symposium on LSP, 1997*, vol. 1. Copenague (Dinamarca), Copenhagen Business School. Págs. 349-359.

TopQuadrant (2010): *TopBraid Composer, Getting Started Guide*, versión 3.0. TopQuadrant [publicación electrónica]. Disponible en línea: <http://www.topquadrant.com/docs/marcom/TBC-Getting-Started-Guide.pdf>

Tsarkov, D. y I. Horrocks (2006): «FaCT++ Description Logic Reasoner: System Description». En *Proceedings of the International Joint Conference on Automated Reasoning (IJCAR 2006)*, vol. 4130 de *Lecture Notes in Artificial Intelligence*, Springer-Verlag. Págs 292-297. Disponible en línea: <http://www.cs.man.ac.uk/~tsarkov/papers/TsHo06a.pdf>

UMLS (2009): *UMLS Reference Manual*. Bethesda (EEUU), National Library of Medicine. Disponible en línea: <http://www.ncbi.nlm.nih.gov/books/NBK9676/>

Van Heijst, G.; A. T. Schreiber y B. J. Wielinga (1997): «Using explicit ontologies in KBS development». En *International Journal of Human and Computer Studies*, vol. 45. Amsterdam (Holanda), Elsevier Science Publishers. Págs. 183-292. Disponible en línea: <http://www.cs.vu.nl/~guus/papers/Heijst97a.pdf>

Vaupel, J. (2005): «Modelling Building Element Classes and Properties». En *Workshop on Modelling Classes & Properties in Terminology & Ontologies, 7th International Conference on Terminology and Knowledge Engineering*, 16 de agosto de 2005, Copenague (Dinamarca), Copenhagen Business School.

Winston, M.; R. Chaffin y D. J. Herrmann (1987): «A taxonomy of Part-Whole Relations». En *Cognitive Science*, vol. 11. Págs. 417-444

Wright, S. E. y L. Wright (1997). «Descriptive Terminology: Terminology Management for Technical Translation». En S.E Wright y G. Budin (eds.): *Handbook of Terminology Management*, vol. 1 Amsterdam, John Benjamins. Págs. 147-159.

Wüster, E. (1998): *Introducción a la teoría general de la terminología y a la lexicografía terminológica*. Barcelona, IULA.

Glosario de acrónimos y abreviaturas

AENOR	Asociación Española de Normalización y Certificación
BCT	base de conocimiento terminológico
BDT	base de datos terminológica
DIN	Deutsches Institut für Normung (Instituto alemán de normalización)
DL	description logics (lógica descriptiva, lógica de descripciones)
KWIC	key word in context (palabra clave en context)
OWL	ontology web language (lenguaje de ontologías web)
TBM	Terminología Basada en Marcos
TCT	Teoría Comunicativa de la Terminología
TGT	Teoría General de la Terminología
TSC	Terminología Socio-Cognitiva

Índice de figuras

Figura 1. Ejemplo de estructura léxica de los términos derivados de <i>relay</i> (relé) (Roche 2009: 2)	30
Figura 2. Ontología de <i>relay</i> (relé) (Roche, 2009: 2).....	30
Figura 3. Extracto de una ontoterminología de turbinas hidráulicas y su término en francés (Roche, 2009: 4).	32
Figura 4. Espacio multidimensional en un sistema de conceptos (Kageura, 1997: 121) ..	41
Figura 5. Sistema de conceptos en CAOS 1 (Madsen et al., 2005: 166)	66
Figura 6. Extracto de la clase ‘contagious-disease’ en Genoma-KB	69
Figura 7. Resultado de la búsqueda de ‘erosión’ en EcoLexicon	71
Figura 8. Recurso auxiliar con imágenes sobre los conceptos en EcoLexicon.....	72
Figura 9. Muestra de la interfaz de acceso a Oncoterm vía web.....	73
Figura 10. Flujo de trabajo en <i>Termtontology</i> (ibídem: 5)	84
Figura 11. Módulo de la ontología	
Figura 12. Ejemplo de pantalla de edición en el módulo de la base de datos	92
Figura 13. Muestra de la interfaz de Ontoterm	92
Figura 14. Interfaz de <i>Termtontology Workbench</i>	94
Figura 15. Arquitectura de un sistema de representación basado en la lógica de descripciones (Baader y Nutt, 2003: 50).....	106
Figura 16. Espectro de ontologías basado de Jimeno-Yepes et al. (Jimeno-Yepes et al., 2008:2)	109
Figura 17. Taxonomía de clases de OWL-KR (Gómez-Pérez et al., 2004: 67).....	111
Figura 18. Extracto de la jerarquía de categorías <i>top-level</i> de Sowa	111
Figura 19. Representación de un extracto de la red semántica de UMLS (UMLS, 2009)	112
Figura 20. Estructura doble para un adjetivo en <i>Wordnet</i>	114
Figura 21. Clase ‘humano’ bajo la clase ‘Thing’	118
Figura 22. Clase ‘humano’ como subclase de ‘mamífero’	118
Figura 23. Clase ‘humano’ y su instancia ‘María’	118
Figura 24. La instancia ‘María’ con la datatype property ‘tieneEdad’ y valor ‘28’	119
Figura 25. <i>Anotation property</i> en la clase ‘humano’ mediante <i>label</i> para indicar denominación equivalente en inglés.	120
Figura 26. Inferencia de la clase ‘María’ en base a que ‘tieneHijo’ es subpropiedad de ‘tienePariente’	121
Figura 27. Inferencia en la instancia ‘Pau’ en base a que la clase ‘humano’ equivale a la clase ‘persona’	121
Figura 28. Inferencia en la instancia ‘Juan’ en base a que ‘tieneFamiliar’ es equivalente a ‘tienePariente’	122
Figura 29. Explicación de porqué la instancia ‘Pau’ no pertenece a la clase ‘mujer’ en base a que es disjunta a ‘hombre’	122
Figura 30. Inferencia de que la instancia ‘Pau’ tiene la propiedad ‘tienePadres’ con valor María en base a su inversa.....	123
Figura 31. Inferencia entre instancias en base a una propiedad transitiva	124
Figura 32. Inferencia entre instancias en base a una propiedad simétrica	124
Figura 33. Inconsistencia detectada en la instancia en base a una propiedad funcional. 124	
Figura 34. Inferencia en base un axioma de equivalencia basado en una restricción de cardinalidad mínima.....	125

Figura 35. Inferencia de la relación entre las clases ‘hombre’ y ‘varón’ basada en la restricción del rango en la propiedad ‘tieneGénero’ que las describe.....	126
Figura 36. Inferencia de la relación entre las clases ‘arroses’ y ‘paella’ en base a la definición del rango de ‘tieneIngrediente’ que las describe.....	127
Figura 37. Inferencia de la relación entre las clases ‘Menorca’ y ‘lugares’ en base a la definición del rango de ‘tieneGenticilicio’ que las describe.	127
Figura 38. Proceso de desarrollo y ciclo de vida de <i>Methontology</i> (ibídem: 127).....	133
Figura 39. Ejemplo de diagrama de relaciones binarias	134
Figura 40. Interfaz de Protégé 4 con la ontología <i>pizza</i> abierta	142
Figura 41. Visualización de la pestaña Entities con la ontología <i>pizza</i> abierta.....	143
Figura 42. Visualización de la pestaña Classes con la ontología <i>pizza</i> abierta.....	144
Figura 43. Visualización de la pestaña Object Properties con la ontología <i>pizza</i> abierta.....	144
Figura 44. Visualización de la pestaña Data Properties con la ontología <i>pizza</i> abierta ..	145
Figura 45. Visualización de la pestaña Individuals con la ontología <i>pizza</i> abierta.....	145
Figura 46. Visualización de la pestaña OntoGraf con el esquema de un fragmento de la ontología <i>pizza</i>	146
Figura 47. Visualización de la pestaña DL Query con una consulta sobre pizzas que contienen anchoas utilizando el razonador HermiT.	146
Figura 48. Creación de clase <i>pizza</i> bajo la clase <i>Thing</i> en la ontología <i>Pizzaejemplo</i>	147
Figura 49. Visualización desde la pestaña <i>Classes</i> de Protégé 4 de un fragmento de la jerarquía de <i>pizza</i>	147
Figura 50. Ejemplo de <i>object property</i> en la ontología <i>Pizzaejemplo</i>	148
Figura 51. Ejemplo de <i>datatype property</i> en <i>Pizzaejemplo</i> y ventana de asignación del rango.....	149
Figura 52. Instancia <i>margarita_Etrusco</i> de la clase <i>margarita</i> en <i>Pizzaejemplo</i>	150
Figura 53. Descripción de la clase <i>margarita</i> en <i>Pizzaejemplo</i>	151
Figura 54. Descripción de la clase <i>vegetariana</i> como clase equivalente en <i>Pizzaejemplo</i>	151
Figura 55. Inferencia de la clase <i>margarita</i> como subclase de <i>vegetariana</i> en la ontología <i>Pizzaejemplo</i>	152
Figura 56. Instancias de la clase <i>restaurante</i> en <i>Pizzaejemplo</i>	153
Figura 57. Aserción en la instancia <i>margarita_Etrusco</i> de la propiedad <i>tiene_restaurante</i> en <i>PizzaEjemplo</i>	153
Figura 58. Aserción en la instancia <i>margarita_Etrusco</i> de <i>tiene_precio</i> y ejemplo de anotación.....	154
Figura 59. Anotación en la propiedad <i>tiene_ingrediente</i> de tipo <i>label</i>	155
Figura 60. Anotación de la propiedad <i>tiene_precio</i> utilizando como tipo <i>comment</i>	155
Figura 61. Anotación, mediante <i>label</i> , de denominaciones para la propiedad <i>tiene_ingrediente</i>	156
Figura 62. Visualización de propiedades en la ontología <i>Pizzaejemplo</i> a través de las etiquetas en idioma inglés.	156
Figura 63. Axioma general en <i>Pizzaejemplo</i>	157
Figura 64. Resultado de la consulta sobre pizzas con origen Toscana en <i>Pizzaejemplo</i>	157
Figura 65. Resultado de la consulta en DL sobre pizzas con jamón.....	158
Figura 66. Interfaz de TopBraid con la ontología <i>pizza ontology</i>	161
Figura 67. Descripción de la clase <i>CheesyPizza</i> en la ventana central	162
Figura 68. Posibilidades de búsqueda desde TopBraid.....	163

Figura 69. Descripción gráfica de baldosa y las distintas medidas que se consideran en la norma (AENOR, 2007:14)	172
Figura 70. Búsqueda de “resistencia” y “abrasión”	180
Figura 71. Creación de clases y subclases desde <i>Create Class Hierarchy</i> en Protégé y ejemplo de estructura desde Word	182
Figura 72. Ejemplo de grafo generado con <i>OWL 2 Viz</i>	182
Figura 73. Ejemplo de <i>class matrix</i> con la propiedad ‘tipo_de_característica’ y tipo de descripción ‘some’	183
Figura 74. Menú de uso del razonador en Protégé 4.....	183
Figura 75. Visualización de información inferida por HermiT.....	184
Figura 76. Ejemplo de error por valor mal expresado	184
Figura 77. Relación entre plenitud de superficie: curvatura central, curvatura lateral y alabeo.	200
Figura 78. Modelo estático de descripción de una característica y su rol en la descripción de conceptos.	205
Figura 79. Descripción de ‘pavimento de bicocción’	217
Figura 80. Grafo del primer nivel de la jerarquía de clases de la ontología.....	226
Figura 81. Grafo de la jerarquía de segundo nivel de la clase ‘cerámica’	226
Figura 82. Formalización de clases disjuntas.....	227
Figura 83. Ejemplo de jerarquía desde la pestaña <i>object property</i>	227
Figura 84. Jerarquía de <i>object property</i> en la ontología.....	228
Figura 85. Descripción de la <i>datatype property</i> para ‘aspecto superficial’	228
Figura 86. Datatype properties que representan características	229
Figura 87. Ejemplo de instancias de la subclase de acabado ‘esmaltado’ y descripción de la instancia ‘alto brillo’ como un tipo de ‘esmaltado’, visible desde la pestaña	229
Figura 88. Ejemplo de <i>annotation properties</i> incluidas en la descripción de la clase <i>pequeñas diferencias de color</i>	230
Figura 89. Descripción de la clase ‘espesor’	231
Figura 90. Vínculo de una propiedad que representa a la característica ‘acabado’ con la clase ‘acabado’	233
Figura 91. Instancias de la clase ‘color’ que representan los valores de ‘color’	234
Figura 92. Valores subclase de ‘acabado’	235
Figura 93. Subvalores del valor ‘esmaltado’	235
Figura 94. Subvalores del valor ‘no esmaltado’	235
Figura 95. Formalización de la característica ‘acabado’ como <i>object property</i>	235
Figura 96. Formalización de la característica ‘parámetro de ensayo’ como <i>object property</i>	236
Figura 97. Formalización de ‘resistencia a la helada’ como <i>datatype property</i>	236
Figura 98. Descripción de dominio y rango de ‘tiene alabeo’	237
Figura 99. Descripción de dominio y rango de ‘tiene resistencia a la abrasión superficial’	238
Figura 100. Extracto de la norma UNE-EN 10545-8.....	239
Figura 101. Ejemplo ficticio de uso de la propiedad de dilatación térmica lineal	239
Figura 102. Anotaciones y descripción del dominio y rango de la propiedad de ‘aspecto superficial’	240
Figura 103. Descripción en las clases ‘capacidad de absorción de agua’ y ‘porosidad aparente’ de la equivalencia entre ellas.....	241

Figura 104. Inferencia en la propiedad ‘tiene capacidad de absorción de agua’ de la propiedad equivalente ‘tiene porosidad aparente’	241
Figura 105. Anotaciones de la propiedad de ‘resistencia a la abrasión superficial’	242
Figura 106. Anotación sobre la fuente de la clase ‘aspecto superficial’ y sobre otra fuente relevante	242
Figura 107. Ejemplo de contextos de ‘aspecto superficial’	243
Figura 108. Anotación en el valor ‘pasta roja’ de la característica ‘composición’	243
Figura 109. Anotación sobre el parámetro ‘dureza de Mohs’	243
Figura 110. Observación sobre los valores posibles de ‘tiene aspecto superficial’ en forma de anotación con el subelemento <i>comment</i>	244
Figura 111. Formalización de la relación baldosa cerámica – proceso de fabricación ...	244
Figura 112. Formalización de la relación característica - proceso	245
Figura 113. Dominio y rango de la propiedad para la relación de meronimia.....	245
Figura 114. Declaración de clases parte de la clase ‘planitud de superficie’	246
Figura 115. Declaración de la superclase de ‘curvatura’ y ‘alabeo’	246
Figura 116. Dominio y rango de la propiedad que representa el tipo de característica ..	247
Figura 117. Descripción de la clase ‘apariencia’ como característica intrínseca	247
Figura 118. Descripción de la equivalencia en la clase ‘característica intrínseca’	248
Figura 119. Descripción de la equivalencia en la clase ‘característica extrínseca’	248
Figura 120. Jerarquía de clases de ‘apariencia’	249
Figura 121. Fragmento de la descripción de ‘acabado’ con la característica distintiva ‘tiene parámetro de ensayo’	250
Figura 122. Extracto de la descripción de ‘aspecto superficial’ con una característica esencial delimitada	251
Figura 123. Extracto de la descripción de la clase ‘apariencia’ con la característica distintiva ‘tiene parámetro de ensayo’	252
Figura 124. Descripción de la clase ‘baldosa esmaltada’ y su característica distintiva ..	252
Figura 125. Descripción de la clase ‘baldosín catalán’ y sus características distintivas.	253
Figura 126. Descripción de la clase de baldosa cerámica ‘grupo A’ con las característica esenciales remarcadas	253
Figura 127. Anotaciones sobre otra denominación y el equivalente en inglés de ‘aspecto superficial’	255
Figura 128. Estructura de clases vista a través de las denominaciones en inglés	255
Figura 129. Consulta sobre las clases subordinadas a la clase ‘resistencia química’ a través de su denominación en inglés	256
Figura 130. Fragmento de la jerarquía de clases de la ontología	256
Figura 131. Formalización de la característica ‘aspecto superficial’ desde la pestaña de <i>Classes</i>	257
Figura 132. Descripciones heredadas de la superclases de ‘aspecto superficial’	258
Figura 133. Formalización de la propiedad que representa a ‘aspecto superficial’ desde la pestaña <i>Data properties</i>	258
Figura 134. Descripción de la clase ‘P’ que representa el parámetro de ensayo de ‘porosidad aparente’	259
Figura 135. Formalización de la relación <i>característica – proceso de ensayo</i>	259
Figura 136. Primer nivel de clases de la ontología	261
Figura 137. Visualización de la herencia múltiple en clases de la cerámica	262
Figura 138. Creación de clases y subclases desde <i>Create Class Hierarchy</i> en Protégé y ejemplo de estructura desde Word	263

Figura 139. Grafo sobre ‘indicadores de absorción de agua’ y relaciones entre conceptos subordinados	264
Figura 140. Valores de textura.....	265
Figura 141. Descripción en lenguaje natural del valor pasta roja en la ontología	265
Figura 142. Descripción en forma de Object property de ‘tiene acabado’	266
Figura 143. Referencia al parámetro de ensayo de ‘tiene acabado’	266
Figura 144. Anotación sobre el vínculo entre la clase ‘acabado’ y la propiedad ‘tiene acabado’	267
Figura 145. Formalización de los valores de ‘acabado’ en la ontología	268
Figura 146. Complemento para la descripción de color	268
Figura 147. Formalización del dominio y rango de ‘pequeñas diferencias de color’	270
Figura 148. Anotaciones en ‘tiene resistencia a la helada’	270
Figura 149. Dominio y rango de ‘tiene alabeo’	271
Figura 150. Restricción de valores posibles en “tiene a la abrasión superficial”	271
Figura 151. Anotaciones de ‘tiene_alabeo’	272
Figura 152. Anotaciones de ‘tiene dilatación por humedad’	272
Figura 153. Dominio y rango de ‘aspecto superficial’	273
Figura 154. Definición de la <i>object property</i> ‘tipo de característica’	274
Figura 155. Definición de ‘acabado’ como característica intrínseca	274
Figura 156. Visualización de extracto de la pestaña <i>Class matrix</i> en Protégé 4.....	275
Figura 157. Definición de la clase ‘extrínseca’	275
Figura 158. Estructura inferida por el razonador	276
Figura 159. Introducción de la relación <i>característica – proceso de ensayo</i> en las características correspondientes	277
Figura 160. Clases de la ontología que tienen la relación característica – proceso de ensayo con algún proceso de ensayo.....	277
Figura 161. Descripción de la clase ‘resistencia a ácidos y álcalis en concentraciones fuertes’	279
Figura 162. Descripción de la clase ‘característica de la baldosa cerámica’	280
Figura 163. Contextos en la clase ‘alabeo’	280
Figura 164. Anotación sobre la norma de referencia de la característica ‘alabeo’.	281
Figura 165. Anotación sobre el catálogo donde podemos encontrar la característica ‘acabado’	281
Figura 166. Clases de grupos UNE de baldosa	282
Figura 167. Clase de baldosa cerámica con característica distintiva ‘tiene capacidad de absorción de agua’	283
Figura 168. Clase de baldosa cerámica que se distingue por el valor en ‘relación baldosa – proceso de fabricación’	283
Figura 169. Clase de baldosa cerámica que se distingue por la combinación de valores en ‘tiene capacidad de absorción de agua’ y ‘relación baldosa – proceso de fabricación’	284
Figura 170. Visualización de la descripción de características esenciales a través de <i>Class Matrix</i>	285
Figura 171. Clases de pavimento y revestimiento según la UNE 14411	285
Figura 172. Clases extraídas del manual de PROALSO.....	285
Figura 173. Comentario sobre los valores según el parámetro ‘dureza de Mohs’	286
Figura 174. Comentario sobre el procedimiento de limpieza A	286
Figura 175. Comentario de la clase ‘moldeo’	287
Figura 176. Introducción de equivalentes y sinónimos en forma de <i>annotation</i>	287

Figura 177. Extracto de la descripción de 'color' y la clase que se ha eliminado de la descripción en la prueba.....	289
Figura 178. Fragmento de la estructura inicial de clases de 'característica de la baldosa cerámica'	290
Figura 179. Extracto de la clasificación de 'características de la baldosa cerámica' después de desagruparlas	290
Figura 180. Grafo de las subclases inferidas de 'característica de la baldosa cerámica'	291
Figura 181. Resultado de la consulta sobre la clase de 'resistencia a las manchas'	292
Figura 182. Extracto del resultado de la consulta sobre valores de 'color'	292
Figura 183. Resultado de la consulta sobre los parámetros de ensayo de 'capacidad de absorción de agua'	293
Figura 184. Resultado de la consulta sobre baldosas con resistencia a la helada	293
Figura 185. Explicación de 'baldosión catalán' como subclase de 'grupo AIIb'	294
Figura 186. Explicación del razonador de la equivalencia de 'baldosín catalán' y 'barro cocido'	294

V. ANEXOS

Anexo 1 – Acceso a la tabla de descripción general de características

La tabla de descripción general de características es un fichero de cálculo de gran extensión, para ser explorada con comodidad, aconsejamos descargar el fichero del siguiente enlace y abrirlo con Microsoft Excel.

Enlace a la tabla: <http://goo.gl/3kVAzO>



QR:

Anexo 2 - Fichas descriptivas de baldosas

Anexo 3 - Listado de instrumentos

Instrumento de ensayo de características
agente de manchas
agente de manchas verde en aceite ligero
agente de manchas rojo en aceite ligero
yodo
aceite de oliva
agente limpiador
agua caliente
agente limpiador de débil actividad
agente limpiador de fuerte actividad
agua
agua desionizada
anillo metálico
aparato de abrasión
aparato de abrasión PEI
aparato frigorífico
aparato para valoración visual
aparato calibrado para medir la dilatación térmica
arnés
autoclave
azul de metileno
balanza
baño para baja temperatura
barra de referencia de acero al níquel
bola de acero cromado
bomba de vacío
calibre
cámara de vacío
cesta
cilindro de vidrio borosilicato
colorímetro triestímulo
comparador
dsecador
detergente
disolvente
acetona
hidróxido potásico
ácido clorhídrico
dispositivo para impregnación con agua
espectrofotómetro de reflectancia
espectrómetro de absorción atómica
estufa

estufa de secado
fotómetro
fuentes de calor
galga de espesor
gamuza
horno
lámpara eléctrica
lanzador de bola
lápiz
luz fluorescente
manómetro registrador
material abrasivo
material para sellar
mecanismo de cronometraje electrónico
mineral
plucómetro
micrómetro
micrómetro con palpadores
micrómetro nonio
pañó blanco
placa patrón
probeta graduada
recipiente con tapadera de vidrio borosilicato
regla
regla metálica
rodillo central cilíndrico
rodillo de apoyo cilíndrico de metal
sellador a base de silicona
solución acuosa
producto de limpieza doméstica
cloruro amónico
sal para piscina
hipoclorito sódico
ácido
ácido acético
ácido clorhídrico
ácido cítrico
ácido láctico
base
hidróxido potásico
tapadera impermeable
termopar
vaso de vidrio

Anexo 4 - Listado de parámetros

nombre del parámetro	contexto
P	La porosidad aparente, P, expresada en porcentaje, es la relación existente entre el volumen de los poros abiertos de la probeta y su volumen exterior. [UNE-EN ISO 10545-3]
cociente masa seca/volumen exterior	
coeficiente de absorción de agua	Para cada baldosa, el coeficiente de absorción de agua, $E_{(b,v)}$, expresado como porcentaje de masa seca, se calcula mediante la ecuación $E_{(b,v)} = [(m_2^{(b,v)} - m_1) / m_1] \times 100$ donde m_1 es la masa de la baldosa seca; m_2 es la masa de la baldosa húmeda. [UNE-EN 10545-3]
coeficiente de dilatación lineal	El coeficiente de dilatación lineal de origen térmico, a_1 , se expresa con aproximación de una cifra decimal $\times 10^{-6}$ por grado centígrado ($10^{-6}/^{\circ}\text{C}$) [UNE-EN 10545-8]
coeficiente de fricción	
coeficiente de restitución	
coeficiente de restitución	
Cromatidad	
	Borde_rugoso
	Burbuja
	cuarteo
	Defecto bajo el esmalte

Defecto de decoración	
Desvitrificación del esmalte	
Falta de esmalte	
Grieta	
Mella	
ondulado	
pinchado	
punto o mancha	
rebaba	
defecto por alta temperatura	
defecto por baja temperatura	
defecto por presión	
cuarteo	
defecto visible a revoluciones	defecto visible a revoluciones según etapas de abrasión clase 0= 100 clase 1= 150 clase 2 = 600 clase 3= 750, 1 500 clase 4= 2 100, 6 000, 12 000 clase 5= >12 000(1) (1) Debe superar el ensayo especificado en ISO 10545-14 para la resistencia a las manchas [UNE-EN 10545-7]
diferencia de color	
diferencia de longitud / Longitud inicial por mil	

dureza de Mohs	1 Talco Se raya fácilmente con la uña 2 Yeso Se raya con la uña 3 Calcita Se raya con moneda de cobre 4 Fluorita Se raya fácilmente con lámina de acero 5 Apatito Se raya dif. con lámina de acero 6 Feldespato Se raya con lima metálica 7 Cuarzo Dif. con lima metálica 8 Topacio Raya el vidrio de ventana 9 Corindón Raya el vidrio de ventana 10 Diamante Raya el vidrio de ventana [CE020-2e.txt]
efecto de decoración	
espacio de color CIELAB	
factor comercial	
fuerza de rotura por grosor	
grado de resistencia a sustancias	
	grado de resistencia a manchas
	grado de resistencia a ácido y álcali
	grado de resistencia a producto de limpieza
	grado de resistencia a sal para piscinas
grosor medio de cada lado	
indicadores de absorción de agua	

Densidad aparente	<p>Contexto 1: Se produce por una irregular distribución de la densidad aparente en la pieza que motiva a su vez una diferencia de contracción lineal entre distintas zonas de la pieza, siendo mucho más acusada en productos de pavimento que presentan una contracción de cocción superior al 5%. [Ce004-2e.txt]</p> <p>Contexto 2: La densidad aparente, B, expresada en gramos por centímetro cúbico, de una muestra, es el cociente entre su masa seca y el volumen exterior, incluidos los poros. [UNE-EN ISO 10545-3]</p>
densidad relativa aparente	<p>Contexto 1: 772 densidad relativa. Cociente de la densidad de una sustancia (un gas, por ej.) por la de otra tomada como patrón (hidrógeno, por ej.), bajo condiciones similares de presión y temperatura. [CED01-1e.txt]</p> <p>Contexto 2: El método de la ebullición debe ser utilizado para la clasificación y la caracterización de las baldosas. El método del vacío debe emplearse para determinar la porosidad abierta, la densidad relativa aparente y la absorción de agua salvo, en este caso, a efectos de clasificación de las baldosas. [UNE-EN ISO 10545-3]</p>

	<p>Contexto 1: Con la definición "Producto gresificado" se entiende todas aquellas piezas esmaltadas o no que después de cocidas presentan una absorción de agua (porosidad aparente) inferior al 6%. [CE005-3.txt]</p> <p>Contexto 2: La porosidad aparente, P, expresada en porcentaje, es la relación existente entre el volumen de los poros abiertos de la probeta y su volumen exterior. [UNE-EN ISO 10545-3]</p> <p>Contexto 3: No se recomienda el uso de ceras, aceites o productos similares, ya que la nula porosidad de la superficie vidriada impide la correcta aplicación. [Cat12.pdf]</p>
porosidad aparente	
lugar de instalación	
luminosidad	
masa impermeable	
material de fabricación	
medida de lados cortos	
medida de lados largos	
porcentaje de baldosas libre de efectos	
porcentaje de desviación	
	porcentaje de desviación de la arista
	porcentaje de desviación de la recta de un lado
	porcentaje de desviación del vértice
	porcentaje de desviación del centro de la baldosa

porcentaje de desviación del ángulo de la arista
solubilidad de la sustancia
tratamiento de la cara vista
volumen de materia eliminada

Anexo 5 - Listado de procesos

Proceso de ensayo de la característica	contexto
ensayo con inmersión	
ensayo de la raya de lápiz	
ensayo de reflexión	
ensayo sin inmersión	
impregnación de agua	
	método al vacío
método PEI	
	método por ebullición
pesaje hidrostático	
procedimiento de limpieza	
	procedimiento de limpieza A
	Limpiar la probeta con agua corriente caliente durante 5 min, después enjuagar con un paño húmedo. [UNE-EN 10545-14]
	procedimiento de limpieza B
	Limpiar la probeta a mano con el agente limpiador de débil actividad, utilizando una esponja natural no abrasiva o un paño. Después enjuagar la superficie con agua corriente, y finalmente enjuagarla con un paño húmedo. [UNE-EN 10545-14]
	procedimiento de limpieza C
	Limpiar la probeta por medios mecánicos con el agente limpiador de fuerte actividad (5.1.3) [UNE-EN 10545-14]

	Sumergir la probeta durante 24 h en un disolvente apropiado (5.1.4), después enjuagar abundantemente con agua corriente, y finalmente, enjuagarla con un paño húmedo. La limpieza se considera terminada si uno de los disolventes (5.1.4) hace desaparecer la mancha [UNE-EN 10545-14]
procedimiento de limpieza D	
recocido	
tratamiento con agua hirviendo	
Procesos de fabricación de la baldosa	
cocción	
	bicocción
	monococción
moldeo	

	<p>Baldosas extruidas (moldeo A) Baldosas cuya masa se moldea en estado plástico mediante una galletera, y la cinta obtenida se corta en piezas de longitud predeterminada. [...] Baldosas prensadas en seco (moldeo B): Baldosas formadas de una masa reducida a polvo o pequeños granos y moldeadas en matrices a alta presión. Pueden ser esmaltadas o no esmaltadas.[...] Baldosas coladas (moldeo C): Baldosas formadas de una masa en estado de barbotina que se vierte en un molde o en una placa refractaria porosa que absorbe el agua. [...]Baldosas coladas (moldeo C): Baldosas formadas de una masa en estado de barbotina que se vierte en un molde o en una placa refractaria porosa que absorbe el agua. [Ce008-1e.txt]baldosas cerámicas: Son placas de poco espesor fabricadas con arcillas y/o otras materias primas inorgánicas, generalmente utilizadas como revestimiento de suelos y paredes, moldeadas por extrusión (A) o por prensado (B) a temperatura ambiente, aunque pueden fabricarse mediante otros procedimientos (C), seguidamente secadas y posteriormente cocidas a temperaturas suficientes para desarrollar las propiedades necesarias. Las baldosas pueden ser esmaltadas (GL) o no esmaltadas (UGL) y son incombustibles e inalterables a la luz. [UNE-EN 14411]El gres porcelánico está comprendido dentro del grupo Bla (Baldosas cerámicas prensadas en seco con absorción de agua E < 0,5%) de la norma ISO 13006 y UNE 67-087. [Cat15.pdf]</p>
<p>extrusión</p>	
<p>otros procesos</p>	
	<p>encolado manual</p>

Baldosas extruidas (moldeo A) Baldosas cuya masa se moldea en estado plástico mediante una galletera, y la cinta obtenida se corta en piezas de longitud predeterminada. [...] Baldosas prensadas en seco (moldeo B): Baldosas formadas de una masa reducida a polvo o pequeños granos y moldeadas en matrices a alta presión. Pueden ser esmaltadas o no esmaltadas.[...] Baldosas coladas (moldeo C): Baldosas formadas de una masa en estado de barbotina que se vierte en un molde o en una placa refractaria porosa que absorbe el agua. [...]Baldosas coladas (moldeo C): Baldosas formadas de una masa en estado de barbotina que se vierte en un molde o en una placa refractaria porosa que absorbe el agua. [Ce008-1e.txt]

baldosas cerámicas: Son placas de poco espesor fabricadas con arcillas y/o otras materias primas inorgánicas, generalmente utilizadas como revestimiento de suelos y paredes, moldeadas por extrusión (A) o por prensado (B) a temperatura ambiente, aunque pueden fabricarse mediante otros procedimientos (C), seguidamente secadas y posteriormente cocidas a temperaturas suficientes para desarrollar las propiedades necesarias. Las baldosas pueden ser esmaltadas (GL) o no esmaltadas (UGL) y son incombustibles e inalterables a la luz. [UNE-EN 14411]

El gres porcelánico está comprendido dentro del grupo Bla (Baldosas cerámicas prensadas en seco con absorción de agua E < 0,5%) de la norma ISO 13006 y UNE 67-087. [Cat15.pdf]

prensado en seco

Anexo 6 - Listado de unidades de medida

centímetro

gramo

gramo por centímetro cúbico

miligramo por decímetro

cuadrado

milímetro

milímetro cúbico

milímetro por metro

Newton por milímetro cuadrado

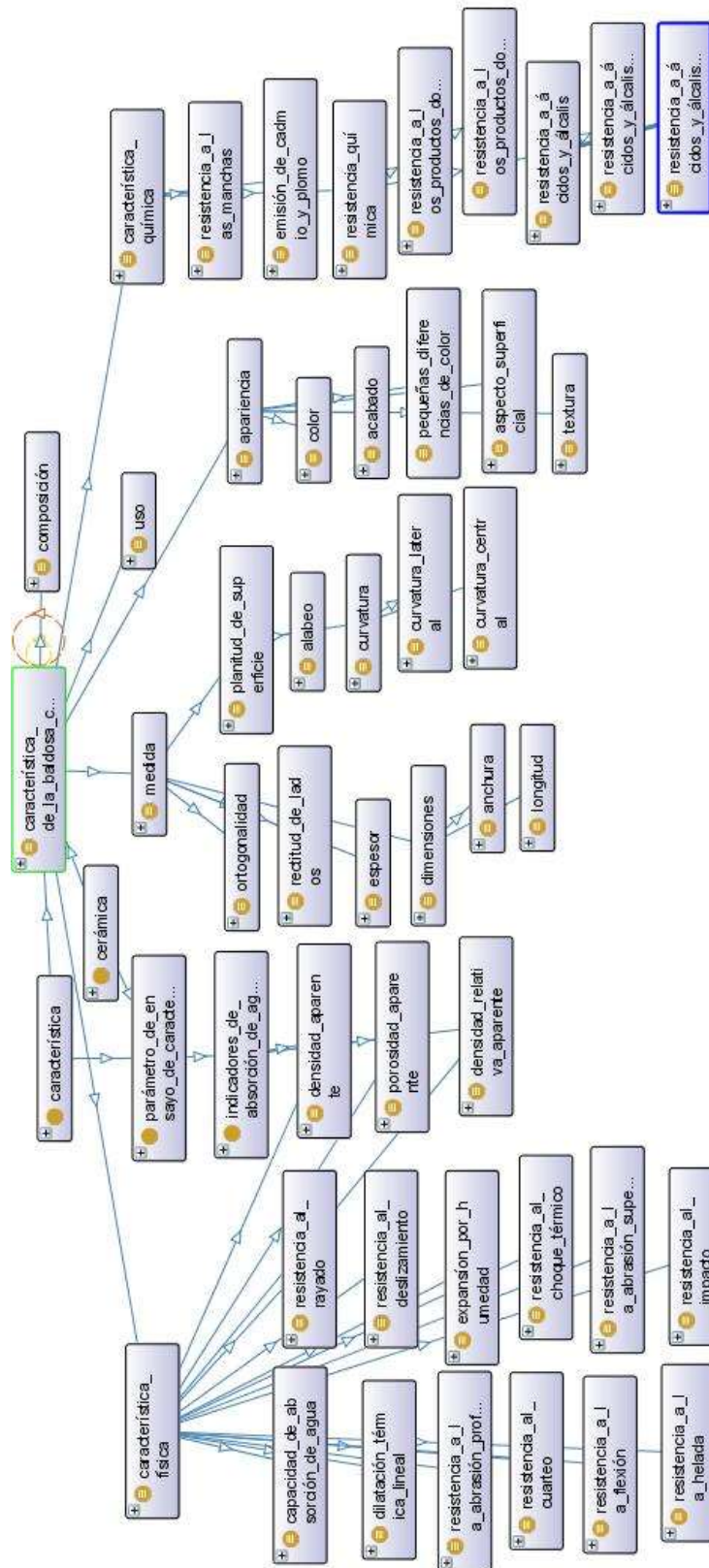
Anexo 7- Glosario de equivalencias en español e inglés de las características y baldosas

Denominación de la característica	Otras denominaciones	Equivalente en inglés
acabado		finish
alabeo		warpage
anchura	ancho	width
aspecto superficial	calidad superficial	surface quality
capacidad de absorción de agua	absorción de agua	water absorption
color		colour
composición		materials
curvatura		centre curvature
curvatura central		edge curvature
curvatura lateral		apparent density
densidad aparente		bulk density
densidad relativa aparente	densidad relativa	moisture expansion
dilatación térmica lineal	expansión térmica	linear thermal expansion
dimensiones	tamaño, formato	dimensions
emisión de cadmio y plomo		lead and cadmium release
espesor	grosor	thickness
expansión por humedad	expansión por humedad	moisture expansion
longitud		length
ortogonalidad		rectangularity
pequeñas diferencias de color		small colour differences
planitud de superficie	planitud	surface flatness
porosidad aparente	porosidad, porosidad abierta	apparent porosity
rectitud de lados		straightness of sides
resistencia a ácidos y álcalis	resistencia a los ácidos y las bases	resistance to chemicals
resistencia a ácidos y álcalis en concentraciones débiles	resistencia a los ácidos y las bases en concentraciones débiles	resistance to low concentrations of acids and alkalis
resistencia a ácidos y álcalis en concentraciones fuertes	resistencia a los ácidos y las bases en concentraciones fuertes	resistance to high concentrations of acids and alkalis
resistencia a la abrasión profunda	resistencia a la abrasión UGL	resistance to deep abrasion
resistencia a la abrasión superficial	resistencia a la abrasión GL	resistance to surface abrasion
resistencia a la flexión	módulo de rotura	modulus of rupture
resistencia a la helada		frost resistance
resistencia a las manchas		resistance to staining

resistencia a los productos domésticos de limpieza	resistencia a los productos de limpieza	resistance to household chemicals
resistencia a los productos domésticos de limpieza y aditivos para agua de piscinas		resistance to household chemicals and swimming pool salts
resistencia al choque térmico		resistance to thermal shock
resistencia al cuarteo		crazing resistance
resistencia al deslizamiento		friction resistance
resistencia al impacto		impact resistance
resistencia al rayado	dureza al rayado superficial	scratch resistance
resistencia química	resistencia al ataque químico	chemical resistance
textura		shape
uso		use

Anexo 8 – Grafo de las superclases de la ontología

Anexo 9 - Grafo de la clase 'característica de la baldosa cerámica'



Anexo 10 - Bibliografía de recursos textuales

1. Normas de estandarización por organismo y número de norma

AENOR (1997-2001): *UNE EN ISO 10545: Baldosas cerámicas*. Madrid, Asociación Española de Normalización y Certificación (AENOR).

AENOR (1997a): *UNE-EN ISO 10545-1: Baldosas cerámicas. Parte 1: Muestreo y criterios de aceptación*. Madrid, Asociación Española de Normalización y Certificación (AENOR).

AENOR (1998a): *UNE-EN ISO 10545-2: Baldosas cerámicas. Parte 2: Determinación de las dimensiones y del aspecto superficial*. Madrid, Asociación Española de Normalización y Certificación (AENOR).

AENOR (1997b): *UNE-EN ISO 10545-3: Baldosas cerámicas. Parte 3: Determinación de la absorción de agua, de la porosidad abierta, de la densidad relativa aparente, y de la densidad aparente*. Madrid, Asociación Española de Normalización y Certificación (AENOR).

AENOR (1997c): *UNE-EN ISO 10545-4: Baldosas cerámicas. Parte 4: Determinación de la resistencia a la flexión y de la carga de rotura*. Madrid, Asociación Española de Normalización y Certificación (AENOR).

AENOR (1998a): *UNE-EN ISO 10545-5: Baldosas cerámicas. Parte 5: Determinación de la resistencia al impacto por medición del coeficiente de restitución*. Madrid, Asociación Española de Normalización y Certificación (AENOR).

AENOR (1998b): *UNE-EN ISO 10545-6: Baldosas cerámicas. Parte 6: Determinación de la resistencia a la abrasión profunda de las baldosas no esmaltadas*. Madrid, Asociación Española de Normalización y Certificación (AENOR).

AENOR (1999a): *UNE-EN ISO 10545-7: Baldosas cerámicas. Parte 7: Determinación de la resistencia a la abrasión superficial de las baldosas esmaltadas*. Madrid, Asociación Española de Normalización y Certificación (AENOR).

AENOR (1997d): *UNE-EN ISO 10545-8: Baldosas cerámicas. Parte 8: Determinación de la dilatación térmica lineal*. Madrid, Asociación Española de Normalización y Certificación (AENOR).

AENOR (1999b): *UNE-EN ISO 10545-9: Baldosas cerámicas. Parte 9: Determinación de la resistencia al choque térmico*. Madrid, Asociación Española de Normalización y Certificación (AENOR).

AENOR (1997e): *UNE-EN ISO 10545-10: Baldosas cerámicas. Parte 10: Determinación de la dilatación por humedad*. Madrid, Asociación Española de Normalización y Certificación (AENOR).

AENOR (1997f): *UNE-EN ISO 10545-11: Baldosas cerámicas. Parte 11: Determinación de la resistencia al cuarteo de baldosas esmaltadas*. Madrid, Asociación Española de Normalización y Certificación (AENOR).

AENOR (1997g): *UNE-EN ISO 10545-12: Baldosas cerámicas. Parte 12: Determinación de la resistencia a la helada*. Madrid, Asociación Española de Normalización y Certificación (AENOR).

AENOR (1998c): *UNE-EN ISO 10545-13: Baldosas cerámicas. Parte 13: Determinación de la resistencia química*. Madrid, Asociación Española de Normalización y Certificación (AENOR).

AENOR (1998d): *UNE-EN ISO 10545-14: Baldosas cerámicas. Parte 14: Determinación de la resistencia a las manchas*. Madrid, Asociación Española de Normalización y Certificación (AENOR).

AENOR (1998e): *UNE-EN ISO 10545-15: Baldosas cerámicas. Parte 15: Determinación de la emisión de plomo y cadmio en las baldosas esmaltadas*. Madrid, Asociación Española de Normalización y Certificación (AENOR).

AENOR (2001): *UNE-EN ISO 10545-16: Baldosas cerámicas. Parte 16: Determinación de pequeñas diferencias de color*. Madrid, Asociación Española de Normalización y Certificación (AENOR).

AENOR (2003): *UNE-ENV 12633: Método para la determinación del valor de la resistencia al deslizamiento/resbalamiento de los pavimentos pulidos y sin pulir*. Madrid, Asociación Española de Normalización y Certificación (AENOR).

AENOR (2007): *UNE 14411: Baldosas cerámicas. Definiciones, clasificación, características y marcado*. Madrid, AENOR. Madrid, Asociación Española de Normalización y Certificación (AENOR).

AENOR (2011): *UNE-EN 15771: Esmaltes vítreos y de porcelana. Determinación de la dureza al rayado de la superficie según la escala de Mohs*. Madrid, Asociación Española de Normalización y Certificación (AENOR).

ISO (1995a): *ISO 10545-1: Ceramic tiles. Part 1: Sampling and basis for acceptance*. Ginebra, International Organization for Standardization (ISO).

ISO (1995b): *ISO 10545-2: Ceramic tiles. Part 2: Determination of dimensions and surface quality*. Ginebra, International Organization for Standardization (ISO).

ISO (1995c): *ISO 10545-3: Ceramic tiles. Part 3: Determination of water absorption, apparent porosity, apparent relative density and bulk density*. Ginebra, International Organization for Standardization (ISO).

ISO (1994a): *ISO 10545-4: Ceramic tiles. Part 4: Determination of modulus of rupture and breaking strength*. Ginebra, International Organization for Standardization (ISO).

ISO (1996a): *ISO 10545-5: Ceramic tiles. Part 5: Determination of impact resistance by measurement of coefficient of restitution*. Ginebra, International Organization for Standardization (ISO).

ISO (1995d): *ISO 10545-6: Ceramic tiles. Part 6: Determination of resistance to deep abrasion for unglazed tiles*. Ginebra, International Organization for Standardization (ISO).

ISO (1996b): *ISO 10545-7: Ceramic tiles. Part 7: Determination of resistance to surface abrasion for glazed tiles*. Ginebra, International Organization for Standardization (ISO).

ISO (1994b): *ISO 10545-8: Ceramic tiles. Part 8: Determination of linear thermal expansion*. Ginebra, International Organization for Standardization (ISO).

ISO (1994c): *ISO 10545-9: Ceramic tiles. Part 9: Determination of resistance to thermal shock*. Ginebra, International Organization for Standardization (ISO).

ISO (1995e): *ISO 10545-10: Ceramic tiles. Part 10: Determination of moisture expansion*. Ginebra, International Organization for Standardization (ISO).

ISO (1994d): *ISO 10545-11: Ceramic tiles. Part 11: Determination of crazing resistance for glazed tiles*. Ginebra, International Organization for Standardization (ISO).

ISO (1995f): *ISO 10545-12: Ceramic tiles. Part 12: Determination of frost resistance*. Ginebra, International Organization for Standardization (ISO).

ISO (1995g): *ISO 10545-13: Ceramic tiles. Part 13: Determination of chemical resistance*. Ginebra, *International Organization for Standardization (ISO)*.

ISO (1995h): *ISO 10545-14: Ceramic tiles. Part 14: Determination of resistance to stains*. Ginebra, *International Organization for Standardization (ISO)*.

ISO (1995i): *ISO 10545-15: Ceramic tiles. Part 15: Determination of lead and cadmium given off by glazed tiles*. Ginebra, *International Organization for Standardization (ISO)*.

ISO (1999): *ISO 10545-16: Ceramic tiles. Part 16: Determination of small colour differences*. Ginebra, *International Organization for Standardization (ISO)*.

ISO (1998): *ISO 13006: Ceramic tiles. Definitions, classification, characteristics and marking*. Ginebra, *International Organization for Standardization (ISO)*. Ginebra, ISO.

2. Catálogos comerciales (con código de referencia en ontología)

APARICI CERÁMICAS (2008a): *ARQ 2008*. Alcora, Aparici Cerámicas S.A. [Cat01.pdf]

APARICI CERÁMICAS (2008b): *Catálogo Cevisama '08*. Alcora, Aparici Cerámicas, S.A. [Cat03.pdf]

APARICI CERÁMICAS (2009a): *Catálogo General '09*. Alcora, Aparici Cerámicas S.A. [Cat04.pdf]

APARICI CERÁMICAS (2009b): *Catálogo Novedades '09*. Alcora, Aparici Cerámicas S.A. [Cat05.pdf]

APARICI CERÁMICAS (2009c): *Exclusive Collection*. Alcora, Aparici Cerámicas S.A. [Cat12.pdf]

AZULEJOS SANCHIS (2009c): *Catálogo General 2009*. Alcora, Azulejos Sanchis S.L. [Cat06.pdf]

HALCÓN CERÁMICAS (2009): *Catálogo general 2009*. Alcora, Halcón Cerámicas, S.A. [Cat14.pdf]

HALCÓN CERÁMICAS [En línea]: *Consejos de mantenimiento del pavimento porcelánico pulido* <http://www.halconceramicas.com/ima/contenidos/consejos%20material%20pulido.pdf> Fecha de consulta: 07 de marzo de 2012. [Cat17.pdf]

HALCÓN CERÁMICAS [En línea]: *Consejos de colocación* <http://www.halconceramicas.com/ima/contenidos/consejos%20de%20colocacion.pdf>. Fecha de consulta: 07 de marzo de 2012. [Cat18.pdf]

HALCÓN CERÁMICAS (2009): *Novedades 2009*. Alcora, Halcón Cerámicas S.A. [Cat13.pdf]

HISPANO AZUL (2009a): *Catálogo general de interior*. Alcora, Hispano Azul S.A. [Cat02.pdf]

HISPANO AZUL (2009b): *Crea 2009*. Alcora, Hispano Azul S.A. [Cat16.pdf]

HISPANO AZUL (2009b): *Porcelánico Alta Gama*. Alcora, Hispano Azul S.A. [Cat15.pdf]

PAMESA CERÁMICA (2008): *Catálogo general 2008*. Pamesa. Castellón, Pamesa Cerámica S.L. [Cat00 (en papel)]

STYLNUL CERÁMICA (2006): *Catálogo general 2006*. Nules, Stylnul Cerámica S.A. [Cat07.pdf]

STYLNUL CERÁMICA (2007): *Catálogo general 2007*. Nules, Stylnul Cerámica, S.A. [Cat08.pdf]

STYLNUL CERÁMICA (2008): *Catálogo general 2008*. Nules, Stylnul Cerámica, S.A. [Cat09.pdf]

STYLNUL CERÁMICA (2009): *Catálogo general 2009*. Nules, Stylnul Cerámica, S.A. [Cat10.pdf]

VIVES AZULEJOS Y GRES (2009): *Catálogo general 2008-2009*. Alcora, Vives Azulejos y Gres, S.A. [Cat11.pdf]

3. Bibliografía de la versión en español del corpus TXTCeram

ALARCÓN, J.; J. CARDA y J. M. RINCÓN (1992): *Nuevos productos y tecnologías de esmaltes y pigmentos cerámicos. Su fabricación y utilización*, Castellón, Faenza Editrice Ibérica, S.L.

AMORÓS, J. L.; A. BARBA y V. BELTRÁN (1994): *Estructuras cristalinas de los silicatos y óxidos de las materias primas cerámicas*, Castellón, Instituto de Tecnología Cerámica-Asociación de Investigación de las Industrias Cerámicas.

AMORÓS, J. L.; V. BELTRÁN; A. BLASCO; J. E. ENRIQUE; A. ESCARDINO y F. NEGRE (1991): *Defectos de fabricación de pavimentos y revestimientos*, Castellón, Asociación de investigación de las industrias cerámicas.

ATC (1990): *Tecnología de la fabricación de azulejos*. Castellón, Asociación Técnicos Cerámicos (ATC).

BARBA, A; V. BELTRÁN y C. FELIU (1997): *Materias primas para la fabricación de soportes de baldosas cerámicas*, Castellón, Instituto de Tecnología Cerámica.

CLAUSELL, A. y R. GALINDO (1991): *Apuntes de operaciones básicas en la fabricación de pavimentos y revestimientos cerámicos*, Castellón, Conselleria de Cultura, Educació i Ciència, Generalitat Valenciana.

ENRIQUE, J. E.; F. NEGRE; V. BELTRÁN y J. GUILLEM (1980): *Mayólica, gres y porcelana artística*, Castellón, Asociación de investigación de las industrias cerámicas.

ESCARDINO, A. y M. GONZÁLEZ-CUDILLEIRO (1991): *Azulejos y pavimentos cerámicos españoles*, Ministerio de Industria, Comercio y Turismo.

ESCRIBANO, P.; J. B. CARDA y E. CORDONCILLO (2001): *Esmaltes y pigmentos cerámicos*, Castellón, Faenza Editrice Iberica.

FERNÁNDEZ-NAVARRO, J. M. (1991): *El vidrio*, Madrid, Consejo Superior de Investigaciones Científicas. 2ª ed. (reimpresión)

GALINDO, R. (1994): *Pastas y vidriados*, Castellón, Faenza Editrice Iberica.

GARCÍA-VERDUCH, A. (1993): *Colocación de pavimentos y revestimientos cerámicos*, Castellón, Instituto de Tecnología Cerámica.

GIOVANNINI, R. (1989): *La serigrafía en la cerámica*, Barcelona, Ediciones Omega, S.A.

LUCAS, F. (2004): *Tensión superficial en suspensión de esmaltes*, *Frittainnova*, 1(1).

GUILLEM-MONZONIS, C. y M. C. GUILLEM-VILLAR, (1987): *Diccionario cerámico científico-práctico (español-inglés-alemán-francés)*. Castellón, Sociedad española de cerámica y vidrio.

MATTHES, W. E. (1990): *Vidriados cerámicos. Fundamentos, propiedades, recetas, métodos.*, Barcelona, Omega.

MONRÓS, G.; J. A. BADENES; A. GARCÍA y M. A. TENA (2003): *El color de la cerámica: Nuevos mecanismos en pigmentos para los nuevos procesados de la industria cerámica*, Castellón, Publicacions de la Universitat Jaume I. Servei de Comunicació y Publicació.

NEBOT-DÍAZ, ISAAC; M. MARCHAL; M. IRÚN y J. B. CARDA (2000): *Nuevas tecnologías para el sector cerámico de Castellón. Desarrollo de esmaltes vitrocrystalinos y vitrocerámicos*, Castellón, Publicacions de la Universitat Jaume I.

NORTON, F.H. (1988): *Cerámica fina. Tecnología y aplicaciones*, Barcelona, Ediciones Omega.

PADOA, L. (1990): *La cocción de productos cerámicos. Con especial referencia a los materiales de revestimiento y de pavimento*, Barcelona, Omega.

PORCAR, J. L.; A. PLEGUEZUELO; F. RENAU y J. M. GOMIS (1987): *Manual-guía técnica de los revestimientos y pavimentos cerámicos*, Castellón, Instituto de Tecnología Cerámica.

QUALICER (1998a): *Conferencias Generales y ponencias*, V Congreso Mundial de la Calidad del Azulejo y del Pavimento Cerámico, vol. I. Castellón, Cámara oficial de Comercio, Industria y Navegación.

QUALICER (1998b): *Conferencias Generales y ponencias*, V Congreso Mundial de la Calidad del Azulejo y del Pavimento Cerámico, vol. II. Castellón, Cámara oficial de Comercio, Industria y Navegación.

QUALICER (2000a): *Conferencias Generales. Ponencias. Posters. Mesa Redonda*, VI Congreso Mundial de la Calidad del Azulejo y del Pavimento Cerámico, vol. I. Castellón, Cámara oficial de Comercio, Industria y Navegación.

QUALICER (2000b): *Conferencias Generales. Ponencias. Posters. Mesa Redonda*, VI Congreso Mundial de la Calidad del Azulejo y del Pavimento Cerámico, vol. II. Castellón, Cámara oficial de Comercio, Industria y Navegación.

QUALICER (2000c): *Conferencias Generales. Ponencias. Posters. Mesa Redonda*, VI Congreso Mundial de la Calidad del Azulejo y del Pavimento Cerámico, vol. III. Castellón, Cámara oficial de Comercio, Industria y Navegación.

QUALICER (2002a): *Conferencias Generales. Ponencias. Posters. Mesa Redonda*, VII Congreso Mundial de la Calidad del Azulejo y del Pavimento Cerámico, vol. I. Castellón, Cámara oficial de Comercio, Industria y Navegación.

QUALICER (2002b): *Conferencias Generales. Ponencias. Posters. Mesa Redonda*, VII Congreso Mundial de la Calidad del Azulejo y del Pavimento Cerámico, vol. II. Castellón, Cámara oficial de Comercio, Industria y Navegación.

QUALICER (2002c): *Conferencias Generales. Ponencias. Posters. Mesa Redonda*, VII Congreso Mundial de la Calidad del Azulejo y del Pavimento Cerámico, vol. III. Castellón, Cámara oficial de Comercio, Industria y Navegación.

QUIMICER (2004): *Quimicer Spherical*, Castellón, Cámara oficial de Comercio, Industria y Navegación.

RADO, P. (1990): *Introducción a la tecnología de la cerámica*, Barcelona, Ediciones Omega.

RHODES, D. (1999): *Hornos para ceramistas*, Barcelona, Ceac.

SACMI (2004): *Tecnología cerámica aplicada*, Vol II, Castellón, Faenza Editrice Iberica.

TORRELLA, E. (1996): *La producción de frío*, Valencia, Universidad Politécnica de Valencia, Servicio de Publicaciones.

4. Manual de baldosas

PROALSO (2011): “Clasificación y selección de la baldosa cerámica”, en *Manual de actualización en materiales y técnicas de colocación de recubrimientos cerámicos. Carnet profesional Alicatador Solador, Nivel 1: Aplicaciones Convencionales*. Castellón, Asociación Profesional de Alicatadores/Soldadores de España (PROALSO).

5. Diccionario cerámico

GUILLEM-MONZONIS, C. y M. C. GUILLEM-VILLAR, (1987): *Diccionario cerámico científico-práctico (español-inglés-alemán-francés)*. Castellón, Sociedad española de cerámica y vidrio.

Anexo 11 – Información para la instalación del software y acceso a la ontología

La aplicación Protégé 4 y el software complementario recomendado para consultar la ontología está disponible en: <http://goo.gl/Q5QDK>



QR:

En esta dirección, el usuario encontrará:

- La aplicación Protégé 4.1 para su descarga e instalación
- La aplicación Graphviz 2.28 de instalación necesaria para el funcionamiento de la pestaña OWLViz de Protégé 4.1
- Un archivo comprimido con la ontología (“ontoceram712_7.zip”).

Para facilitar la instalación y acceso a la ontología se pone a disposición del lector un vídeo demostración en: <http://goo.gl/0uKg3>. En este vídeo se explica, además, cómo instalar el plug-in *Matrix para Protégé 4.1*. Este plug-in no es necesario para consultar la ontología pero se ha utilizado en la creación y descripción de entidades de la ontología.



QR:

INTRODUCCIÓN Y CONCLUSIONES EN ESPAÑOL

I. Introducción

1 Marco de trabajo

El volumen de información con el que se trabaja en la sociedad actual es el resultado, por una parte, de un desarrollo extraordinario de los medios y recursos para llevar a cabo el intercambio de información y, por otra, de la necesidad de ese intercambio. Siguiendo esta línea, este trabajo parte de la idea de que la reutilización de información y la creación de bases de conocimiento sirven a la sociedad de la información en la gestión eficiente de sus recursos.

Dentro de la investigación sobre el lenguaje orientada a cubrir las necesidades de la sociedad de la información, se encuentra el estudio orientado a cubrir aquellas que puedan surgir a partir de la especialización de los dominios de trabajo. Esta investigación sobre los lenguajes de especialidad encuentra, en un contexto de intercambio de datos, objetos de estudio relacionados con el uso de la terminología propia de una especialidad sobre un determinado dominio con el fin de garantizar el éxito en la comunicación.

Con un enfoque interdisciplinario donde participa de otras disciplinas como la informática, la documentación o la lingüística, la terminología intenta proporcionar el soporte necesario para mejorar el intercambio de información a nivel global. Autores como Temmerman (2000), Faber (Faber, y otros, 2005) o Roche (2005) aportan nuevos enfoques para afrontar las nuevas necesidades del terminólogo y los retos que se le plantean a la terminología moderna en la sociedad de la información. La sistematización de recursos terminológicos o la automatización de tareas relacionadas con la terminología se convierten en factores de gran importancia para la mejora del funcionamiento y del intercambio de la información (Alcina, 2001).

En lo que respecta a las tecnologías desarrolladas al efecto, los avances en investigación abarcan ya muchos campos aplicados. Procesos como la adquisición manual de datos terminológicos a partir de texto se están sustituyendo por técnicas automáticas que ayudan a realizar esta ardua tarea (Feliu, 2004). Además, las bases de conocimiento y sistemas de ontologías se presentan como potentes herramientas conceptuales para gestionar datos terminológicos y conocimiento (Gruber, 1995; Guarino, 1995). En este sentido, despierta nuestro interés la posibilidad de utilizar razonadores de ontologías. Los razonadores de ontologías son programas que permiten inferir información nueva, no representada explícitamente en una ontología, a partir de la información representada (Bock y otros, 2008). En general, en la mayoría de los casos, encontramos que los proyectos adoptan enfoques interdisciplinarios; y entre estas disciplinas se encuentran la terminología y la ingeniería del conocimiento. Actualmente, en la investigación sobre ontologías y terminología, el énfasis se sitúa en la mejora de las herramientas que puedan necesitar aquellos que trabajan con el lenguaje y la información (enciclopedias electrónicas, sistemas de gestión terminológica, aplicaciones de traducción asistida, etc.).

Consciente de que es necesario investigar en sistemas que faciliten el acceso a la terminología especializada y aceleren la creación de recursos destinados a facilitar la

comunicación entre especialistas, este trabajo se propone contribuir a través de la investigación en terminología sobre representación y formalización de características del concepto en una ontología. El marco de este trabajo son los proyectos TxtCeram¹ y Ontodic².

TxtCeram (Alcina, 2006) es un proyecto de investigación que ha sido desarrollado por el grupo de investigación Tecnolettra (Universidad Jaime I). El proyecto se ha centrado en la extracción semiautomática y el análisis conceptual formal de términos de la cerámica. Entre los objetivos del proyecto estaba probar la eficacia de algunas de las herramientas informáticas que se utilizan para diseñar un sistema integral de terminología asistida y los beneficios del uso de este sistema en la mediación lingüística; así como probar la edición de ontologías y estudiar su aplicación para la generación de bases de conocimiento.

Ontodic (Alcina, 2008) es, en parte, la continuación de TxtCeram y ha tenido como objetivo elaborar una metodología sistemática para la elaboración de diccionarios terminológicos onomasiológicos, es decir, diccionarios que contengan terminología especializada en un ámbito de conocimiento y que permitan al usuario consultas a partir del significado, y no solo a partir de la denominación. Los usuarios a quienes va dirigido este tipo de diccionarios son los mediadores lingüísticos (traductores, intérpretes, redactores técnicos), aunque también puede interesar a expertos y al público en general.

El trabajo se centra en la formalización de las características del concepto en forma de ontología. Desde un punto de vista terminológico, los conceptos se pueden clasificar en cuatro tipos de categorías: las entidades, las actividades, las relaciones y las características (Sager y Kageura, 1994: 193). Las características constituyen un elemento de estudio fundamental para describir conceptos y sirven para distinguir unos de otros, y pueden ser muy variadas. Así, encontramos características de distinta naturaleza, como, por ejemplo, en el campo de la cerámica, características de color, de forma o de función. Además, no todas ellas son igual de importantes para la definición del concepto; por ejemplo, encontramos características que sirven para diferenciarlo y otras que son compartidas por todos los tipos de conceptos de un campo. Por esta razón, la descripción de conceptos a partir de la especificación de sus características requiere la indicación del "rango", de la "importancia", de la "naturaleza" o procedencia de cada característica. En este trabajo nos proponemos profundizar en el tipo de información que proporciona una característica y que esta requiere como elemento descriptor y como concepto en sí mismo, observaremos cómo contribuye a la formalización de otros conceptos y cómo podemos formalizarla y representarla en una ontología de manera que proporcione información sobre ella y sobre los conceptos a los que suele describir en un determinado ámbito. Como material de trabajo, nos ocupamos de las características del ámbito de la cerámica, y más concretamente, de la baldosa cerámica.

¹ «TxtCeram: Extracción semiautomática y análisis conceptual formal de términos de la cerámica a partir de un corpus electrónico», proyecto de investigación financiado por la Generalitat Valenciana (código GV05/260). Dirección: Amparo Alcina.

² «ONTODIC I: Metodología y tecnologías para la elaboración de diccionarios onomasiológicos basados en ontologías. Recursos terminológicos para la e-traducción», proyecto de investigación financiado por el Ministerio de Educación y Ciencia, Gobierno de España (código TS12006-01911) y «ONTODIC II: Metodología y técnicas para elaborar diccionarios de colocaciones basados en ontologías. Recursos terminológicos para la e-traducción», proyecto de investigación financiado por el Ministerio de Ciencia e Innovación, Gobierno de España (código TIN2009-07690, subprograma TSI). Dirección: Amparo Alcina.

2 Hipótesis y objetivos

La hipótesis de la cual partimos es que las características de un dominio de especialidad, en este caso el de la ‘baldosa cerámica’, se pueden describir en función de los criterios de clasificación de características con que contamos en los ámbitos de la terminología y la lexicografía, y que esta información conceptual se puede formalizar para representar formalmente las características en una ontología.

El objetivo principal es representar las características de un dominio de especialidad en un sistema de ontologías con un lenguaje de representación de ontologías estándar y teniendo en cuenta los criterios de clasificación de características que aparecen en la bibliografía sobre terminología, lexicografía e ingeniería del conocimiento. El resultado de la representación está orientado a servir de herramienta de consulta a terminólogos y especialistas lingüísticos/técnicos interesados en conocer esta forma de representación de conceptos y, en concreto, la representación de las características de la baldosa cerámica y su naturaleza.

Los objetivos específicos son:

- Identificación de las características de la baldosa cerámica y recopilación de su información conceptual utilizando un enfoque basado en el conocimiento;
- Recogida de una tipología de baldosas utilizando un enfoque basado en el texto;
- Construcción de un modelo de formalización de las características;
- Diseño de una metodología de representación de las características en un editor de ontologías;
- Prueba de la inferencia conceptual aplicada a la representación de características;
- Comparación del enfoque utilizado para representar las características con el enfoque utilizado para representar los tipos de baldosa.

3 Metodología

Uno de los principales retos al abordar este trabajo ha sido encontrar un equilibrio entre el enfoque terminológico del trabajo, el lado técnico que supone el desarrollo de una ontología y el ámbito de la cerámica que nos ocupa, todo sin recaer en demasiados tecnicismos ni teorizaciones.

Para abordar la descripción de las características de la baldosa cerámica hemos realizado un análisis de los criterios de clasificación y descripción presentes en la bibliografía sobre Terminología. Estos criterios y las aportaciones encontradas se han complementado con otras aportaciones más generales sobre descripción de conceptos en Ingeniería del Conocimiento y otras disciplinas.

A continuación, hemos recogido información sobre las características de la baldosa cerámica atendiendo a los resultados del análisis anterior y a las peculiaridades del ámbito. Para ello, hemos extraído información de distintos materiales textuales sobre cerámica y baldosa cerámica y hemos consultado a un experto del ámbito. Durante esta segunda fase de trabajo hemos extraído también equivalentes en inglés de las denominaciones de las características y hemos recogido una tipología de baldosas utilizando un enfoque basado en el texto.

La primera y segunda fase se han llevado a cabo de manera simultánea, ya que a partir del análisis de la bibliografía se han identificado los elementos que forman parte de la descripción de una característica y al mismo tiempo durante la fase de extracción y recogida de datos sobre las características de la baldosa cerámica, la definición de los elementos que forman parte de la descripción de una característica ha adquirido nuevos matices no detectados en la fase de análisis. En el caso de la tipología de baldosas hemos utilizado un enfoque diferente al de las características de la baldosa. Hemos recogido las clasificaciones de tipos de baldosa tal y como figuraban en la bibliografía, sin añadir detalle o alterar sus clases con información de otras fuentes. Como resultado de estas dos fases hemos confeccionado una serie de tablas y fichas que incluyen los elementos que forman parte de la descripción de una característica. Estas tablas y fichas de trabajo se basan en los elementos identificados durante el análisis y en el caso de las características de la baldosa cerámica, en su cotejamiento durante la fase de extracción de información conceptual. Se trata de una compilación de todos los elementos y aspectos a considerar en la creación de la ontología y de las relaciones que se mantienen entre los distintos elementos. El resultado de estas dos fases cumple con los siguientes objetivos:

- Identificación de las características de la baldosa cerámica y recopilación de su información conceptual utilizando un enfoque basado en el conocimiento;
- Recogida de una tipología de baldosas utilizando un enfoque basado en el texto

La construcción de un modelo de formalización de las características constituye el siguiente objetivo del trabajo. La metodología utilizada se ha basado en analizar, desde un punto de vista técnico, la información que queríamos representar en la ontología, los tipos de dato y las relaciones entre estos tipos de datos, y proponer un esquema de formalización válido. Nuestro interés ha sido ser fieles a la información que queríamos representar y mantener un equilibrio entre nuestro objetivo y las posibilidades técnicas de la tecnología utilizada. Esta fase se ha revisado distintas veces a medida que se iba probando la implementación de los datos con la ayuda de un razonador de ontologías.

El trabajo culmina con la implementación de los datos obtenidos en la extracción siguiendo el sistema de formalización elaborado en la fase anterior. Para ello, hemos introducido manualmente y mediante técnicas de inserción automática los datos que configuran la ontología de características de la baldosa cerámica. El modelo de formalización y la implementación de los datos nos han servido para establecer y refinar una metodología de representación de las características en un editor de ontologías. Finalmente, hemos utilizado un razonador para comprobar la consistencia de la ontología y hemos planteado un conjunto de preguntas básicas a las que podría responder la ontología a través de consultas lógicas. Además, hemos comparado la consistencia de la clasificación de características y la consistencia de la clasificación de los tipos de baldosa. El objetivo de la comparación es verificar si la ontología resulta suficientemente consistente, si se ha omitido información relevante para el ámbito y qué

tipo de información es. Hemos analizado los resultados y hemos desarrollado las conclusiones pertinentes en lo que refiere a la metodología y los fundamentos teóricos. El resultado prueba los beneficios del uso de la inferencia conceptual aplicada a la representación de características y nos ha permitido comparar, en base a los datos inferidos, el enfoque utilizado para representar las características con el enfoque utilizado para representar los tipos de baldosa

4 Estructura del trabajo

El presente trabajo se estructura en cuatro grandes bloques: un primer bloque de introducción al trabajo realizado, un segundo bloque donde presentamos los fundamentos teóricos que sirven como base para desarrollar esta tesis, un tercer bloque donde se presenta la descripción del trabajo empírico y finalmente, un cuarto bloque donde presentamos las conclusiones de la tesis doctoral.

En el segundo bloque (§II), analizamos los distintos enfoques teóricos que pueden servir para fundamentar el trabajo empírico realizado y para ello nos centramos, principalmente, en la bibliografía del campo de la Terminología. En este bloque también nos acercamos a la Ingeniería del Conocimiento con el fin de aprovechar algunos de sus planteamientos y tecnologías disponibles.

En el capítulo 1 del bloque de fundamentos teóricos empezamos analizando distintas teorías de la Terminología que sirven de inspiración para llevar a cabo una ontología de características de la baldosa. A continuación, en el capítulo 2 (§II.2), nos centramos en el concepto de características y analizamos su significado a través de distintos autores, introducimos otros conceptos necesarios para entender el significado de característica como *término* o *intensión* para luego adentrarnos de lleno en las características, explorando sus tipos y sus relaciones con otros conceptos. En el capítulo 3 (§II.3) estudiamos el papel de las características en la representación de conceptos, exploramos formas de representar conceptos. El capítulo 4 (§II.4) explora algunas de las metodologías disponibles en la bibliografía sobre Terminología. Finalmente, el capítulo 5 (§II.5) describe algunas herramientas utilizadas en proyectos de creación de ontologías.

Respecto a la ingeniería del conocimiento, en el capítulo 6 (§II.6) iniciamos el recorrido con una breve introducción de los conceptos clave de esta disciplina para terminar adentrándonos en las ontologías, donde presentamos algunos ejemplos de ontologías. En el capítulo 7 (§II.7) exploramos algunos aspectos básicos sobre el lenguaje de ontologías Ontology Web Language 2 (OWL2) y sus posibilidades. En el capítulo 8 (§II.8) presentamos una metodología de elaboración de ontologías y finalmente, en el capítulo 9 (§II.9) nos centramos en las herramientas de edición y otros recursos disponibles para la elaboración de ontologías desde la ingeniería del conocimiento.

El tercer bloque, la descripción del trabajo empírico, se inicia con una introducción al trabajo empírico (§III.1), la descripción de los recursos textuales (§III.2) y de las herramientas (§III.3) utilizadas para llevar a cabo el trabajo. El capítulo 4 (§III.4) describe el análisis de los rasgos que sirven para describir y clasificar las características del concepto a través de distintos trabajos sobre Terminología, Lingüística e Ingeniería del Conocimiento. Al final de este capítulo se presenta el resultado del análisis, que consta de un listado y descripción de elementos que aparecen en la descripción y clasificación de características.

A continuación, en el capítulo 5 (§III.5) describimos el proceso de extracción de características de la baldosa cerámica y su información conceptual y mostramos cómo hemos recogido esta información considerando criterios de descripción, clasificación y validación. Asimismo, describimos el proceso de extracción de la tipología de baldosas cerámicas y sus diferencias con respecto al proceso de extracción de información sobre las características.

En el siguiente capítulo (§III.6) abordamos el diseño de un modelo de formalización de las características en la ontología, este apartado presenta un modelo de formalización de los elementos que sirven para describir una característica en el lenguaje OWL 2 utilizando el editor Protégé OWL 4.1, sus especificaciones y posibilidades. Igualmente, analizamos las distintas formas de representar cada elemento y sus relaciones, las novedades detectadas al aplicar la sintaxis del lenguaje OWL y planteamos una propuesta de formalización acorde con las necesidades del trabajo. En el capítulo 7 (§III.7) presentamos la metodología utilizada para llevar a cabo la introducción de datos en la ontología y se muestran los problemas encontrados y las soluciones que proponemos. Finalmente, en el capítulo 8, analizamos la consistencia en la formalización de características de la baldosa cerámica y la consistencia de la formalización de los tipos de baldosa. Este capítulo también incluye algunos ejemplos de consultas que se pueden realizar mediante un razonador en la ontología. Las conclusiones generales de la tesis se han recogido en el bloque IV.

IV. CONCLUSIONES GENERALES

En este trabajo, hemos logrado formalizar las características de la baldosa cerámica en una ontología y hemos estudiado distintos enfoques de formalización y representación de conocimiento.

En el transcurso del trabajo para construir la ontología de características de la baldosa cerámica, hemos explorado muchas vías de representación y hemos representado muchos tipos diferentes de datos. La ontología constituye un trabajo exhaustivo que requiere un tiempo y dedicación que no siempre está al alcance de otro tipo de proyectos aplicados a fines comerciales. Asimismo, somos conscientes de que puede haber elementos que no hayamos considerado, de forma deliberada o no deliberada, pero estamos seguras de haber creado una ontología cuyo análisis ofrece un gran número de estrategias de representación, de opciones para resolver conflictos y que toma en consideración un amplio abanico de tipos de dato. Consideramos que esto contribuirá al trabajo de futuros desarrolladores de ontologías terminológicas en su tarea de determinar la estrategia a seguir en la representación de según qué tipo de datos, el modelo de formalización adecuado para según qué necesidades o sencillamente, aportar observaciones didácticas sobre cómo transferir conocimiento terminológico a una ontología.

Conclusiones sobre los enfoques y aspectos terminológicos

Los términos se pueden estudiar desde diferentes planos: el lingüístico, el cognitivo y el comunicativo. En este trabajo, en el caso de las características de la baldosa, la extracción de información ha partido de los textos para llegar hasta el conocimiento y se ha combinado con el uso de un experto. En el caso de las baldosas, hemos partido de las denominaciones que hemos identificado en los textos y las hemos descrito siguiendo únicamente los criterios que hemos identificado en los textos.

Creemos que la utilización de las normas de estandarización de la baldosa cerámica y de sus características ha contribuido positivamente a los resultados de este trabajo. Este tipo de documentos están elaborados por expertos y van dirigidos a expertos y tienen la finalidad de organizar y definir el conocimiento con el que trabajan por lo que creemos, que en general, se pueden comparar con los esquemas que realizaría un experto cuando describe un concepto. La combinación de las normas con otro tipo de fuentes nos ha permitido reflejar otras visiones dentro del mismo ámbito e identificar casos de variación de denominaciones, por ejemplo, las denominaciones de los tipos de baldosa o las diferencias entre tipos de característica que hemos tenido que verificar con el experto. El experto ha contribuido a clarificar el conocimiento obtenido, a aportar conocimiento nuevo y a resolver las ambigüedades encontradas.

Hemos probado las diferencias entre lo que consideramos un enfoque semasiológico, tal y como lo concibe la Ontoterminología, donde hemos utilizado únicamente la información representada en los textos para formalizar las baldosas. Sin hacer un análisis posterior del conocimiento subyacente ni contrastar la tipología con expertos del ámbito. Y un enfoque híbrido, con un gran componente onomasiológico, que combina el acercamiento a los conceptos desde su dimensión lingüística pero también desde su dimensión conceptual, ya que no hemos utilizado estos textos únicamente para extraer terminología sino que a través de su lectura y comprensión nos hemos acercado a los conceptos del dominio y al conocimiento que

representan. Y además, hemos recurrido a un experto para validar el conocimiento que queríamos representar.

Al mismo tiempo, hemos incluido en la ontología otra información lingüística y pragmática a través de las anotaciones.

Las características

Hemos formalizado las características en su doble rol de una forma similar a la que plantean enfoques como el de Madsen o Nuopponen: como elementos que describen los rasgos de un concepto y como elementos que constituyen conceptos en sí mismos. Este enfoque defiende que las características que pertenecen a un ámbito de especialidad sean recogidas en los trabajos terminológicos en su doble rol.

Hemos demostrado los conceptos se puede formalizar teniendo en cuenta el criterio de diferencia específica y a la vez dejando lugar a la multidimensionalidad. Esto implica la descripción de múltiples diferencias específicas en casos en los que el concepto tiene más de una dimensión en el dominio específico. Por ejemplo, en el caso de ‘baldosa de gres esmaltada’ existen dos diferencias específicas que no ocurren simultáneamente y dependen de los conceptos coordinados del sistema, por un lado, es necesario representar la diferencia entre baldosa de gres esmaltada y baldosa de gres rústica, donde ésta última no es esmaltada pero también diferenciar entre la baldosa de gres esmaltada y el azulejo esmaltado, donde éste último no contiene gres.

Una definición de característica adecuada para nuestra propuesta es la siguiente: una característica es un concepto cuyo referente es una propiedad de otro concepto.

En la distinción entre relaciones y características coincidimos en parte con el enfoque de Madsen y otros (§2.1.4 y 2.1.5). Desde nuestro punto de vista, en una ontología la representación de una relación y una característica no difieren en forma. La distinción se realiza en la etiqueta que utilicemos para denominar el atributo que representa la característica o la relación según consideremos. Por ejemplo, el dominio y el rango de un atributo que representa la relación proceso-paciente será ‘proceso’ y ‘paciente’ (respectivamente), igualmente, el dominio y rango de un atributo que representa la característica ‘tiene paciente’ en un proceso sería ‘proceso’ y ‘paciente’ (respectivamente). En resumen, consideramos que la representación de la relación entre dos conceptos o de la característica depende de los objetivos de representación concretos, de nuestro interés por reflejar el matiz de relación o de cualidad. Por ejemplo, en la ontología hemos representado la relación “baldosa-proceso de fabricación” en forma de relación, dejando lugar a la extensión de la rama de procesos de esta ontología en futuros trabajos.

La formalización

En lo que se refiere a la metodología de elaboración de una ontología terminológica coincidimos con Roche y otros autores en que es necesario distinguir entre la fase donde se crea un modelo de conocimiento basado en los fundamentos teóricos de la terminología y la fase en la que este modelo se traslada mediante un lenguaje de formalización a una sistema informático que sirve para representarlo. Y para conseguir este objetivo, hemos diferenciado entre la etapa en la que estructuramos, identificamos y delimitamos los elementos necesarios para describir las

características de la baldosa y la etapa en la que diseñamos el modelo que refleja estos elementos en forma de ontología.

Hemos propuesto una metodología o estrategia para realizar la formalización de manera que concepto y término se traten de forma independiente. El objetivo de enfatizar la dimensión conceptual se ha conseguido tomando como punto de partida la estructuración de conceptos, y más concretamente, de las características de las baldosas. En la ontología se distingue una estructura que representa conceptos (con sus características y relaciones) y hace uso de términos normalizados para expresarlos en forma de lenguaje y los términos de uso de cada lengua (a través de la etiqueta “label”). De este modo, el resultado permite que los términos estén relacionados con cada concepto pero sean independientes de la estructura de la ontología.

Los resultados de la prueba de inferencia de las ‘características de la baldosa’ demuestran que la extracción de datos ha sido suficientemente exhaustiva, que la clasificación inicial es consistente y que la formalización de la información es adecuada y se corresponde con el conocimiento que queríamos representar.

Además, la ontología respeta y cumple con todos los principios teóricos planteados en el proceso de formalización:

- Distinción entre característica como concepto y característica como elemento descriptor, formalizada mediante estructuras relacionadas.
- Delimitación de características basadas en la diferencia única y flexibilidad para permitir que las características o las baldosas se agrupen en clases distintas.
- Independencia del lenguaje. Permite adscribir denominaciones en otras lenguas o sinónimos sin afectar a la estructura
- Descripción formal consistente de las características. La descripción de las características tiene el detalle suficiente y necesario como para desagrupar los conceptos de sus clases y que éstos se vuelvan a agrupar utilizando un razonador a través del análisis de características y relaciones en común y distintivas.

En cuanto a los enfoques de formalización y representación de las ‘características de la baldosa’ y los tipos de ‘baldosa cerámica’ consideramos que pueden contribuir como ejemplos dos tipos de enfoque similares a los que distingue la Ontoterminología. En la formalización y representación de características de la baldosa se ha tomado un enfoque que parte de un conocimiento no formal hacia la representación formal del conocimiento. En la formalización y representación de baldosas el enfoque parte de un conjunto de datos y persigue la generación de conocimiento. Es decir, no se representa conocimiento en su clasificación inicial sino que se utilizan las características formalizadas en combinación con el razonador para generar conocimiento a partir de la clasificación inicial.

Las descripciones exhaustivas y consensuadas así como las clasificaciones consistentes utilizadas en la formalización y representación de las características de la baldosa cerámica constituían *per se* conocimiento, y este conocimiento se ha formalizado a través de una ontología.

En cambio, la estructura de tipos de baldosa, parte de una serie de datos y descripciones basadas exclusivamente en la información encontrada en los textos, sin ningún análisis o

interpretación posterior. Su representación en la ontología persigue la obtención de conocimiento a partir del razonamiento y la formalización de esos datos y la información que proporcionan los textos.

Basta con comparar la clasificación inicial de ‘características de la baldosa cerámica’ con la clasificación de ‘baldosas cerámicas’ y estas clasificaciones con las clasificaciones inferidas por el razonador para notar que existen diferencias relevantes entre la representación de ‘características de la baldosa’ y la representación de ‘baldosas cerámicas’. Igualmente, se detectan diferencias en el proceso de formalización e introducción de descripciones. Por ejemplo, la clasificación inicial de ‘características de la baldosa’ permite la distinción entre hermanas de clase basada en la diferencia única. Una única característica o unas mismas características bastan para distinguir las hermanas de una clase. Esto nos sugiere que el espacio de concepto de estas clases está a un mismo nivel. Igualmente, una vez formalizada la descripción de las clases, el razonador ha mantenido la jerarquía de clases y ninguna descripción se ha considerado inconsistente para pertenecer a la clase inicial.

En cambio, la descripción de ‘baldosas cerámicas’ no ha permitido la distinción entre clases por ‘diferencia única’, lo que a nuestro parecer es un síntoma de que la clasificación inicial era inconsistente. Esto se debe a que en la fase de extracción nuestra motivación no era la descripción exhaustiva de tipos de baldosa sino la recogida de ejemplares para luego clasificarlos. Es decir, partíamos de lo expresado en los textos y perseguíamos el conocimiento subyacente en estas descripciones, no presente de manera explícita en los textos.

En lo que refiere al primer enfoque, los resultados son óptimos. Se ha demostrado se pueden describir formalmente las características de la baldosa y que las descripciones que hemos realizado son consistentes y exhaustivas.

En lo que refiere al segundo enfoque, los resultados son alentadores. Hemos probado que el uso de razonadores en la clasificación de conceptos puede ser de gran ayuda para identificar clasificaciones y descripciones poco consistentes. Al mismo tiempo, hemos probado que el uso de características formalmente consistentes para describir conceptos poco consistentes, ayuda a aumentar la consistencia de estos conceptos. Esto se observa por ejemplo en el resultado de clasificación de la clase ‘grupo BIII’ que hemos comentado anteriormente, donde en la descripción que hemos extraído de los textos no se menciona se caracteriza por incluir ‘baldosas esmaltadas’ y no acepta ‘baldosas no esmaltadas’. La utilización en la descripción de la clase de características formalizadas de una manera consistente con dominios y rangos restringidos ha permitido identificar esta característica inherente al concepto que no se había explicitado en los textos.

En lo que refiere a los formalismos, consideramos que en lo que respecta a la formalización de las características esenciales y distintivas todavía queda un gran camino por recorrer. En esta tesis doctoral hemos explorado dos posibilidades en la formalización de características esenciales, una en la que “obligábamos” los miembros de una clase a tener declara la característica esencial y otra en la que permitíamos que una clase se compusiera de miembros que no incumpliendo las restricciones de la clase no tuvieran declarada alguna de las propiedades esenciales.

Consideramos que el primer formalismo sólo es viable si la descripción es consistente y tiene el nivel de detalle suficiente y necesario. Dicho de otro modo, si los conceptos a representar no están bien definidos y las estructuras de conceptos bien delimitadas, la

formalización de las características esenciales restringiendo la necesidad de tener la característica es inviable.

Por eso, recomendamos la formalización utilizada en las características de la baldosa cerámica para características esenciales y no esenciales en los casos en los que el ámbito a representar en la ontología esté bien definido previamente y la formalización utilizada en los tipos de baldosa para los casos en los que no se disponga de toda la información sobre un campo y se quiera aprovechar el razonamiento lógico de una ontología para inferir nueva información.

En cuanto a la contribución de este trabajo a la Terminología, consideramos que permite probar la convivencia interdisciplinar de la Terminología y las posibilidades de las ontologías para la Terminología. En concreto, sirve para contribuir a la creación de ontologías terminológicas que consideren la definición de otro tipo de conceptos más allá de las entidades y reivindica las posibilidades de las características en este tipo de trabajos. En definitiva, esta propuesta pretende contribuir a aumentar las posibilidades de descripción y detalle de las ontologías terminológicas.

Reutilización y duplicación de la investigación

Dejando a un lado las posibilidades de reutilización de la ontología per se en el ámbito de la cerámica, esperamos que este trabajo pueda contribuir como modelo, guía o ejemplo de representación de datos terminológicos en una ontología.

A largo plazo, esta ontología, debido al uso de formatos estándar, se podría utilizar en combinación con otros sistemas de conocimiento, como la web semántica o gestores de conocimiento.

Por otra parte, consideramos que la estrategia para garantizar la homogeneidad y claridad en la elaboración de esquemas conceptuales depende sustancialmente del dominio a tratar. Creemos que en los dominios con una mayor densidad de conceptos que hacen referencia a entidades abstractas y no físicas, la claridad de los esquemas conceptuales resultará más difícil de representar que en el caso de conceptos que hacen referencia a entidades físicas o características medibles.

Igualmente, nos ha llamado la atención que, en la ontología, la característica distintiva de características de la baldosa más frecuente es el parámetro de ensayo, una característica extrínseca. Esto despierta nuestro interés por seguir indagando en este aspecto en otros ámbitos.

Publicaciones relacionadas

Estellés, A. (2013): «Representación de terminología multilingüe en una ontología: un enfoque independiente del concepto», en *New Research in Translation and Interpreting Studies 2013*, Julio, Universitat Rovira i Virgili, Tarragona (España).

Estellés, A. (2012): «Prospects for Building an Ontology on Ceramic Tiles based on ISO Standards information», en *Toth 2012, Terminology and Ontology: Theories and Applications*, Junio, Annecy (France). Disponible en <http://www.porphyre.org/toth/actes>

Estellés, A. y A. Alcina (2012): «Ceramic Tile ISO Standards in the methodology of ontology building», en *Terminology and Knowledge Engineering Conference 2012*, Junio, Madrid (España).

- Estellés, A. y A. Alcina (2009): «A model for formalizing characteristics in Protégé-OWL», en *Proceedings of the 8th International Conference on Terminology and Artificial Intelligence*, Noviembre, Toulouse (Francia). Disponible en <http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-578/>
- Estellés, A. (2008): «Descripción y análisis terminológico de las características del producto cerámico acabado», en *Proceedings of the I Congreso Internacional sobre el Lenguaje de la Ciencia y la Tecnología*, October, Instituto Universitario de Lenguas Modernas Aplicadas (IULMA), Universitat Jaume I, Castellón (España).
- Estellés, A. (2007): «Concept characteristics in the creation of an ontology», en *Workshop Journée thématique sur la terminologie*, Abril, Université de Montréal, Montreal (Canada).
- Estelles, A.; A. Alcina y V. Soler (2006): «Retrieving Terminological Data from the TxtCeram Tagged Domain Corpus: First Step on a Terminological Ontology». En *Proceedings of the International Conference on Language Resources and Evaluation*, May, Genoa (Italy). Págs. 1496-1501. Disponible en <http://www.lrec-conf.org/proceedings/lrec2006/>
- Alcina, A. y A. Estellés (2006): «Etiquetado del corpus TxtCeram orientado a la extracción de información conceptual», en *Proceedings of XXIV Congreso internacional AESLA: Aprendizaje de lenguas, uso del lenguaje y modelación cognitiva: perspectivas aplicadas entre disciplinas*, Abril 2006, UNED, Madrid (España). Págs. 1415-1424