AGENT BASED VIRTUAL ELECTRONIC PATIENT RECORD

From Intra to Inter-Institution Data Integration

PEDRO MANUEL VIEIRA MARQUES

Departament d'Enginyeria de la Informació i de les Comunicacions

Universitat Autònoma de Barcelona

Departament d'Enginyeria de la Informació i de les Comunicacions
Universitat Autónoma de Barcelona

AGENT BASED VIRTUAL ELECTRONIC PATIENT RECORD

From Intra to Inter-Institution Data Integration

PEDRO MANUEL VIEIRA MARQUES

September 2014

Dr. Sergi Robles
(Advisor)

Dr. Ricardo Correia
(External Advisor)

DECLARATION

I certify that I have read this thesis entitled ”Agent Based Virtual Electronic Patient Record - From Intra to Inter-Institution Data Integration” and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Bellaterra, September 2014

Dr. Sergi Robles
(Advisor)

Dr. Ricardo Correia
(External Advisor)

Committee:
Dr. Joan Borrell
Dr. Antonio Moreno
Dr. John McGrory
Dr. Jordi Herrera (substitute)
Dr. Ilídio Oliveira (substitute)

Program:	Doctor en Informàtica.
Department:	Departament d'Enginyeria de la Informació i de les Comunicacions.
Author:	Pedro Manuel Vieira Marques.

To my parents and to Benedita.
Not only in science we stand on the shoulders of giants.

ABSTRACT

Through the years, Health Information Systems (HIS) have been developed and deployed following specific agendas, addressing individual or departamental problems. Albeit several efforts, system integration enhancements are still needed in order to surmount data availability barriers particularly when the status quo reveals that most of the time they coexist as autistic systems. Achieving an integrated and transversal view of all records of one patient is not an easy task as the patterns of data production and usage in healthcare are highly complex, involving heterogeneous actors and an intricate data flow. This is a real issue both within and between health institutions making the integration task difficult and quite commonly not even possible.

Agents are autonomous software entities which can perceive the dynamic character of the surrounding environment enabling pro-activeness regarding the actions that are better suited to a particular user and a given set of goals. They act on behalf of their users and by being socially active they can engage the user, other agents and the environment through message exchanging or auxiliary devices.

In this sense, and by acting on behalf of health professionals in their quest for information, a Multi-Agent approach presents itself as strong candidate for tackling problems in Health Information Systems integration.

The work on this thesis grows from these premises, and is focused on the following questions: Can agent technology enhance or help Health Information Systems integration within a single health institution? How can a single institution agent based approach model be extended for multi-institution health systems integration?

In order to address these questions, a set of objectives were defined: identify the state of the art regarding the use of agents in Healthcare; to address health information integration issues within a single health institution by proposing a model, specification and implementation for agent based intra-institution health data integration; and to address health information integration between different health institutions by extending the models and specification of the previous model to a multiple health institution data integration scenario.

The main contributions from this thesis are: a characterisation of how agent technology is being used for solving problems in the healthcare domain; an agent based system for intra-institution health data integration; a characterisation of health professionals data needs profile; identification of paths for system optimisation and priority management based on type and source of data; and an agent based system for inter-institution health data integration.

RESUM

A través dels anys els sistemes d'informació mèdica (SIM) s'han desenvolupat i desplegat, seguint agendes específiques abordant els problemes individuals. Encara que hi ha hagut diversos esforços, encara és necessari millorar la integració de sistemes per tal de superar les barreres de disponibilitat de les dades, sobretot quan l'*status quo* revela que la majoria de les vegades els sistemes coexisteixen com autistes.

L'assoliment d'una visió integrada i transversal de tots els registres d'un pacient no és una tasca fàcil, ja que els patrons de producció i la utilització de les dades en l'assistència sanitària són molt complexes i involucren actors heterogenis i un flux de dades complexe. Aquest és un problema real, tant dins com entre les institucions de salut que fan la tasca d'integració difícil i molt sovint ni tan sols possible.

Els agents són entitats de programari autònomes que poden percebre el caràcter dinàmic de l'entorn, permetent proactivitat respecte a les accions que s'adapten millor a un usuari particular per a un determinat conjunt d'objectius. Els agents actuen en representació dels seus usuaris, i a través de la seva activitat social poden interactuar amb l'usuari, amb altres agents i amb el propi entorn a través de l'intercanvi de missatges o fent ús de dispositius auxiliars. La seva flexibilitat per integrar altres tecnologies pot millorar l'escalabilitat del sistema i la seva tolerància a fallades, en la direcció d'una inter-operabilitat global dels sistemes.

En aquest sentit, un enfocament multi-agent es presenta com un candidat fort per fer front als problemes de la integració dels Sistemes d'Informació Mèdica.

El treball d'aquesta tesi neix d'aquestes premisses, i es centra en les següents preguntes: pot la tecnologia dels agents millorar o ajudar en la integració de sistemes d'Informació Mèdica dins d'una sola institució sanitària? Com es pot estendre el model basat en agents per a una única institució cap a una integració de sistemes de salut multi-institució?

Per tal d'abordar aquesta qüestió es van definir una sèrie d'objectius: Identificar l'estat de l'art pel que fa a la utilització d'agents en el sector sanitari, adreçar els problemes d'integració d'informació mèdica en una mateixa institució de salut, proposar un model, especificar l'aplicació d'agents per a la integració intra-institució de les dades i informació mèdica entre diferents institucions de salut mitjançant l'ampliació dels models i especificacions del model anterior a un escenari d'integració de dades entre múltiples institucions.

Les principals contribucions d'aquesta tesi són: una caracterització de com s'està utilitzant la tecnologia d'agents per resoldre problemes en l'àmbit mèdic; un sistema d'agents intra-institucional d'integració de dades mèdiques; una caracterització de perfil de dades necessàries de professionals de la salut; la identificació dels camins per a la optimització del sistema i la gestió de prioritats basada en el tipus i la font de les dades; i un sistema basat en agents per a la integració de dades mèdiques inter-institucionals.

RESUMO

Através dos anos, os Sistemas de Informação em Saúde (SIS) foram desenvolvidos seguindo agendas específicas, muitas vezes direcionados para a resolução de problemas particulares e de âmbito departamental. Apesar dos esforços desenvolvidos, a realidade mostra ser ainda necessário o desenvolvimento de processos de integração de sistemas tendo em vista superar as barreiras ainda existentes e permitir a disponibilidade de dados de uma forma transversal. Esta situação é particularmente relevante quando o status quo revela que os sistemas desenvolvidos coexistem de um modo autista. Alterar este estado não é uma tarefa fácil uma vez que os padrões de produção e uso de dados em cuidados de saúde são complexos e envolvem atores heterogêneos. Este é um problema real não só dentro das instituições mas também entre instituições de saúde o que torna a tarefa de integração difícil.

Os Agentes são entidades de software autónomas capazes de se adaptar a um ambiente dinâmico, sendo pró-ativos na identificação das ações que são mais adequadas tendo em vista um conjunto de objectivos pré-definidos. Os Agentes são socialmente ativos e podem interagir com o utilizador, outros agentes e com o seu ambiente através de troca de mensagens ou fazendo uso de dispositivos auxiliares.

Neste sentido, agindo em nome do profissional de saúde na sua busca por informação, a abordagem baseada em Agents apresenta-se como uma forma interessante de mitigar os problemas que dificultam a troca de dados entre sistemas de Informação na Saúde.

O trabalho desta tese evolui a partir destas premissas, e foca-se nas seguintes perguntas: Pode a tecnologia baseada em Agentes ajudar a integração de Sistemas de Informação em Saúde dentro de uma única instituição de saúde? Como pode um modelo baseado em Agentes de integração de dados dentro de uma instituição ser estendido para um ambiente de integração entre sistemas de informação de várias instituições?

Por forma a responder a estas questões, um conjunto de objetivos foram definidos: Identificar o estado da arte em relação ao uso de agentes em na área da Saúde; tendo em consideração os problemas na integração de informação dentro de uma instituição de saúde, propor um modelo, especificação e implementação de um sistema baseado em agentes que promova o acesso alargado aos dados existentes; e alargando a abrangência do acesso aos dados para fora da barreira institucional, estender o modelo e especificações para um cenário de integração de dados entre múltiplas instituições de saúde.

As principais contribuições da tese são: a caracterização de como os Agents são usados na resolução de problemas na área da saúde; um sistema baseado em agentes para a integração de dados de saúde intra-instituição; uma análise de diferentes perfis de profissionais de saúde reflectindo as diferentes necessidades de informação em diversos cenários de prestação de cuidados; a identificação de caminhos para a otimização do sistema e agendamento de ações com base no tipo e fonte de dados; e um sistema baseado em agentes para a integração de dados de saúde entre múltiplas instituições de saúde.

ACKNOWLEDGEMENTS

To Sergi, my UAB supervisor, I would like to acknowledge the acceptance of the challenge of supervising this endeavour and on every occasion encouraging it with an "always look on the bright side of life" word of support.

To Ricardo, my FMUP supervisor, I would like to acknowledge for his long lasting friendship, patience and efforts on pushing this work forward within the labirinthic realm of health information systems.

To the MAID team members Ernesto Palhares, Ana Margarida, Pedro Costa and the SAHIB team members Hilário Almeida, Gustavo Bacelar, Samuel Frade, Cátia Pereira, I would also like to acknowledge their efforts and work, without them this work would not be possible.

To Cristina Santos, I would like to acknowledge for the discussions around the statistical approaches and to Jorge Jácome, for the word of support and availability to discuss speed and acceleration.

To Miguel Angelo, I would like to acknowledge for bearing with my chromatic adventures try-outs.

To all the dEIC Department people, I would like to acknowledge the refreshing environment and the easy-going, friendly immersion into the Catalan way, it will always be a pleasure to return.

At last, my very special thanks to Professor Altamiro for his guidance, safety nets and inspiring thrive, that always made me push forward the limits of reach.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LISTINGS

ACRONYMS

AQL Archetype Query Language

CEN European Committee for Standardisation

CDA Clinical Document Architecture

CPR Computerised Patient Record

CDTE Complementary Diagnosis and Therapeutic Exams

DIS Departamental Information System

EHR Electronic Health Record

ELA Event Listener Agent

EMR Electronic Medical Record

EHR Electronic Health Record

EPR Electronic Patient Record

FIPA Foundation for Intelligent Physical Agents

HIS Health Information System

HL7 Health Level 7

ICD International Classification of Diseases

IGIF Instituto de Gestão Informática e Financeira da Saúde

INE National Institute of Statistics

LBA Local Broker Agent

MAID MultiAgent system for the Integration of Data

MAS Multiagent System

MCA Mobile Collector Agent

MD Medical Doctor

NUTS Nomenclature of Territorial Units

PHR Personal Health Record

RBA Remote Broker Agent

RIM Reference Information Model

RMA Request Manager Agent

RSA Request Scheduler Agent

SAHIB Secure Agent based Health Information systems Brokering

WHO World Health Organisation

VEPR Virtual Electronic Patient Record

Part I

PACKING UP

Log Entry #1 : After defining the starting point, the planned stages and the expected destination of this journey we pack things up.

1 INTRODUCTION

1.1 MOTIVATION

Through the years Health Information Systems (HIS) have been developed and deployed within health institutions following particular agendas, addressing individual or departmental problems. Albeit several efforts, systems integration enhancements are still needed in order to surmount data availability barriers particularly when the status quo reveals that most of the time they result from the coexistence as autistic systems.

However, achieving an integrated and transversal view of all records for a single patient is not an easy task as the patterns of data production and usage in healthcare are highly complex, involving heterogeneous actors and an intricate data flow. These difficulties are a real issue both within and between health institutions making the integration task difficult and quite commonly not even possible.

Agents are software entities, which can embody different perspectives of the surrounding environment and act accordingly. They can perceive the dynamic character of the environment and update their knowledge, enabling proactiveness regarding actions that are better suited according to a particular user and a given set of goals. Agents can act autonomously without the direct action or supervision of users and are socially active through message exchanging with other agents, systems or user. Additionally, agent's flexibility to integrate other technologies may address integration issues, scalability and fault tolerance.

In this sense, and by acting on behalf of health professionals in their quest for information, a Multi-Agent based approach presents itself as a strong candidate for developing systems aimed at tackling problems in Health Information integration.

1.2 OBJECTIVES

The work on this thesis grows from the premises presented above, and is focused on the following questions :

- Can agent technology enhance/help Health Information Systems integration within a single Health Institution?
- How can a single institution agent based approach model be extended for multiple institutions' Health Information Systems integration?

In order to address these questions a set of objectives were defined:

1. Profiling the use of Agent technology in Healthcare.
2. Address Health Information Systems integration issues within a single Health Institution using agent technology by proposing a model, specification and implementation for agent based intra-institution health data integration.

3. Address multiple institutions' Health Information Systems integration issues using agent technology by extending the models and specification to an inter-institution health data integration scenario.

1.3 CONTRIBUTIONS

Following the first objective and making use of systematic review techniques, a large number of projects were identified, reviewed and characterised following a set of classifiers with the purpose of clarifying how agent technology has been being used in Healthcare and aiming at finding relevant gaps that should be address for making these systems adequate to the healthcare environment requirements.

In the following stage, and addressing health data integration issues in the real world scenario, it was designed and developed a model for intra-institutional heals data integration system. The developed system was deployed in a major hospital in Portugal. Further work resulted in a method for evaluating the system execution and enabling system optimisation approaches testing.

In order to understand the scenario of of information sharing between institutions, a characterisation of patient mobility and health institutions usage profile was made along with a characterisation of data usage profiles amongst health professionals providing further insight into the variability of data types used in different care scenarios thus providing hints on how to prevent the harmful effects of delivering a tremendous amount of information to the health professional.

Finally, and pursuing the extension of the previous intra-institutional data integration model, a new model and specification was developed. From this model a pilot system for secure, inter-institutional health data integration was designed, implemented and tested in the laboratory.

PUBLICATIONS

- Cruz-Correia, R.; **Vieira-Marques, P.**; Costa, P.; Ferreira, A.; Oliveira-Palhares, E.; Araujo, F. and Costa-Pereira, A. Integration of hospital data using agent technologies - A case study. Ai Communications, 2005, 18, 191-200[1]
- **P. Vieira-Marques**, R. Cruz-Correia, P. Costa, E. Palhares, A. Ferreira, and A. Costa-Pereira. Maid - Multi Agent for the Integration of Data. Actas Da 1a Conferencia Iberica De Sistemas e Tecnologias De Informacao, Vol I, pages 603–614, 2006.[2]
- **P. M. Vieira-Marques**, R. J. Cruz-Correia, S. Robles, J. Cucurull, G. Navarro, and R. Marti. Secure integration of distributed medical data using mobile agents. Ieee Intelligent Systems, 21(6):47–54, 2006.[3]
- **P. Vieira-Marques**, A. Cunha, L. Antunes, R. Cruz-Correia, and A. Costa-Pereira. A first approach for a regional wide VEPR. HEALTHINF 2008: Proceedings of the 1st International Conference on Health Informatics, Vol 1, pages 215–218, 2008.[4]
- J.H. Patriarca-Almeida, **P.M. Vieira-Marques**, and R.J. Cruz-Correia. Simulation environment for the optimization of the data retrieval capabilities of

an agent based system in a healthcare setting. Holonic and Multi-Agent Systems for Manufacturing Lecture Notes in Computer Science Volume 6867, 2011, pp 124-132 .[5]

- G. M. Bacelar-Silva, **P. M. Vieira-Marques**, and R. J. Cruz-Correia. Identifying relevant data along selected medical care situation. 6th Iberian Conference on Information Systems and Technologies (CISTI), Vol I, pages 350–353, 2011.[6]
- J.H. Patriarca-Almeida, **P.M. Vieira-Marques**, and R.J. Cruz-Correia. Optimization of an agent based clinical data retrieval system. In Information Systems and Technologies (CISTI), 2012 7th Iberian Conference on, pages 1–6, June 2012.[7]
- Silva-Ferreira, P.; Patriarca-Almeida, J.; **Vieira-Marques, P.** ; Cruz-Correia, R. Improving expressiveness of agents using openEHR to retrieve multi-institutional health data: Feeding local repositories through HL7 based providers Information. , 2012 7th Iberian Conference on Systems and Technologies (CISTI), 2012, 1-5[8]
- **P. Vieira-Marques**, R. J. Cruz-Correia, and Sergi Robles, Profiling Roles and Scenarios of Agent application in Healthcare - a Systematic Review. VIII Workshop on Agents Applied in Health Care, held in conjunction with the 14th Conference on Artificial Intelligence in Medicine (AIME 2013)[9]
- **Vieira-Marques, P.**; Patriarca-Almeida, J.; Frade, S.; Bacelar-Silva, G.; Robles, S. ; Cruz-Correia, R. OpenEHR aware multi agent system for inter-institutional health data integration Information , 2014 9th Iberian Conference on Systems and Technologies (CISTI), 2014[10]

DEPLOYED SYSTEMS

- MAID System

PROTOTYPES

- MAID System
- MAID Simulation Environment System
- SAHIB System

OTHER PUBLICATIONS

- R. Cruz-Correia, **P. Vieira-Marques**, A. Ferreira, E. Oliveira- Palhares, P. Costa, and A. Costa-Pereira. Monitoring the integration of hospital information systems: how it may ensure and improve the quality of data. Medical and Care Compunetics 3, 121:176–182, 2006. Studies in Health Technology and Informatics [11].
- R. J. Cruz-Correia, **P. M. Vieira-Marques**, A. M. Ferreira, F. C. Almeida, J. C. Wyatt, and A. M. Costa-Pereira. Reviewing the integration of patient data: how systems are evolving in practice to meet patient needs. BMC Medical Informatics and Decision Making, 7, 2007 [12].

- Ferraz, V.; Oliveira, G.; **Vieira-Marques, P.** ; Cruz-Correia, R. Organs transplantation - how to improve the process? Studies in Health Technologies and Information, 2011, 169, 300-4 [13].

MSC STUDENTS SUPERVISION

- J. H. Patriarca-Almeida - “Optimization of an Agent Based Clinical Data Retrieval System”.
- P. Roberto-Ferreira “Enabling agents to retrieve openEHR-based health data through implementing HL7 communication with departmental information systems”.

PARTICIPATION IN RELATED FUNDED PROJECTS

- Project ICU - a Virtual Electronic Patient Record system development contract for Hospital S. João - 2003
- Project SAHIB - Enhancing multi-institution health data availability through multi-agent systems - 2009/13
- Project HCA - Healthcare Anywhere - 2010
- Project N-KHRONOUS - Security in mobile environments with intermittent connectivity for their medical and aerospace application (2011/14)

PRIZES

- “e-Saúde: Hospital do Futuro” in 2005
- "Fernandes da Costa” in 2006

1.4 ROADMAP

The first part of this thesis consists on a chapter regarding thesis’ motivation and objectives; a chapter where the main concepts of Health Information Systems and Multi-Agent Systems (MAS) technology are described providing some examples of their articulation; finally a chapter where, by applying systematic reviews technics to study the literature, a characterisation of the use of agents in healthcare is presented.

On the second part, a first chapter will describe the intra institution agent based system and its implementation results.

In the following chapter we will be firstly looking into how patients use health services by describing patterns of mobility and secondly characterise health professionals data needs profile by analysing the variability attributed to data in different contexts of healthcare.

Next, the inter-institutional agent based system is described and its implementation results are presented.

The last part will provide a discussion about the journey taken and a final wrapping with suggestions for future work.

2 BACKGROUND

In the next sections we will travel trough Healthcare Information Systems' concepts and take a look into the issues contributing for the complexity underlying the development of information systems that address the need for health data exchange. We will then proceed to describe Agents and Multi-Agent Systems concepts and to explore how these can be used to address Health Information Systems integration issues.

2.1 PATIENT RECORDS AND BEYOND

"The clinical case record freezes in time lifelong events, it is a story in which patient and family are the main characters, with the doctor serving a dual purpose as both biographer and part of the plot. The content of this biography varies considerably, reflecting its many purposes: to recall observations, to inform others, to instruct students, to gain knowledge, to monitor performance and to justify interventions" [14].

The traditional medical record structure originates from the beginning of the twentieth century. In this time additional diagnostic procedures, like radiology and pathology, became important and the hospital grew out to be the cornerstone of medical care. Initially, each department in the hospital kept its own medical record, but in 1907 it was first recommended that all departmental records should be putted together and grouped into one single clinical record per patient [15].

What is an Electronic Health Record?

Trough time medical records have become richer and more complete, aggregating the lifelong events of patient health. As complementary diagnosis means evolve and become more available, the medical data produced increases in size and complexity and so did the medical record supporting it.

Paper was, and in some cases still is, the main support for these records. Although some advantages from paper records are still pointed out (user familiarity hence no need to acquire new skills, portability, flexibility in recording data), the disadvantages regarding content organisation, reduced availability and lack of internal linkage amongst records from a patient outgrew them [16].

Not surprisingly, the paper approach could not keep the pace with the evolution of the healthcare practice and management's evolving requirements [16] and along with the advent and widespread of computers a shift towards making medical reports electronic grew stronger and led to several efforts for developing electronic representations of the traditional paper patient record [17][18].

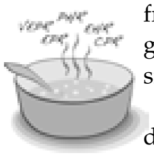
Along with other designations, the Computerised Patient Record (CPR) was generically used to refer to any system that captures, stores, manages or transmits information related to the health of individuals or the activities of organisations that work within the health sector (a term that is now somehow being dropped out of use).

2.1.1 *Electronic Health Records*

As the CPR evolved new designations emerged with the objective of creating wrappers for a particular set of common characteristics regarding data origin, organisation, users and purpose.

The Electronic Medical Record (EMR) and the Electronic Patient Record (EPR) can be seen as a single institution and in some cases, single department hillness oriented system where patient illnesses events solely are recorded. They are used to manage patient data within a single institution. Although some authors have extended the boundary of EPR outside the institutional borders.

The Electronic Health Record (EHR) is a record that manages patient data from multiple departments in multiple institution. It includes information regarding patient during episodes of care provided by different health care professionals [19] and location.



The soup of records

EHRs were classified on the basis of the International Organization for Standardization (ISO) definition [20]. According to this definition, the EHR means a "repository of patient data in digital form, stored and exchanged securely, and accessible by multiple authorised users. It can be seen as an aggregation of several EPR in a single system available throughout the place of care.

A Virtual Electronic Patient Record (VEPR) is virtual in that "it is a view of the data that might be configured differently at different locations, but that is mapped into a common format at the time the record is required" [21]. The VEPR can be situated along the way between an EPR and an EHR. Patient information must appear to the user as a unified set of data even though it may be spread all over the country. The user's view might access only a specially tailored subset of the record in order to handle issues of displaying the information in an intelligible manner [21]. In [22] van Der Linden defines a virtual electronic health record as "an EHR that contains all primarily medical, information on a patient, stored in a variety of systems in a variety of locations over a long period of time, secured against illegal access, provided with an audit trail, and presented to the reader as one dossier." Additionally, other authors state that a virtual electronic patient record based on pre-existing information systems could help the integration process and facilitate the communication among them preventing loss of existent data or without interfering with future software developments [23].

In the meanwhile, with the growing importance attributed to preventive medicine and continuous care not only events of illnesses and their treatments are considered necessary to be stored but also the additional data resulting from events that may influence people health (e. g. eating habits, family history). In this context, a new flavour for electronic health records appeared having a particular focus on patient empowerment regarding its health data management. The Personal Health Record (PHR) is an individual record owned by the patient and that allows him to be more actively involved in its own healthcare by providing the patient with ways of inserting and sharing its health data. Although the designation includes the word *patient*, it goes further beyond the illness condition that the term *patient* implies addressing the individual's lifestyle, habits and additional health related issues.

While EHR systems focus on information needs of health care professionals, PHR systems address the patient point of view capturing health data entered by individuals and provide information closer to the care of those individuals [24].

Notwithstanding, PHR could also be put under the umbrella of an EHR resulting in a more complete and richer data set to be made available at all moments of life to all active participants in the care process.

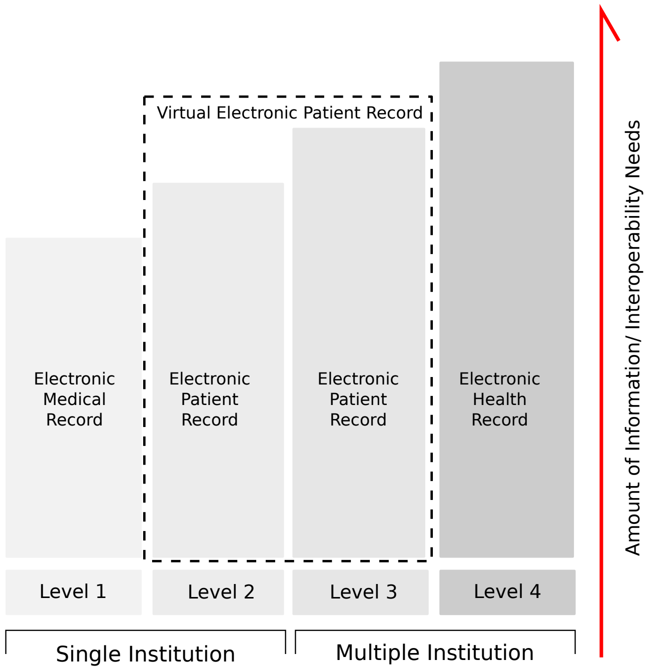


Figure 1: Patient Records classification according to institutional reach and amount of information/interoperability needs.

In this work, it is adopted the hierarchy depicted in Figure 1 regarding the institutional reach, the levels of information (ranging from a single institution, single department information system to a wide range holistic system) and the interoperability needs that grow as the amount, the complexity of data, and the number of interveners increases.

2.1.2 *Advantages of having integrated Electronic Health Records*

Why is HIS integration important ?

Health care is driven by information and knowledge, and its quality depends on taking decisions at the right time and place, according to the right patient data and applicable knowledge hence the need for integration is reinforced. Good medical practice support may be improved by having correct information, available where and when is needed [25]. This requisites cross all borders of care provision from the medical practice to the administrative and management processes.

Nowadays most patients receive care from many health care providers, and consequently their health data are dispersed over many institutions' paper and electronic based record systems. This reality leads to a fragmented system of storing and retrieving essential patient data which may deter patients from having access to optimal care [24].

Changing this status may have a strong impact regarding patient safety, as more information is available to clinicians and new data validation procedures maybe implemented [26] promoting care quality and reducing medical errors. This can lead to reductions in mortality, complications, and costs [27]. Also, by making available increasing amounts of medical information that can be used beyond their original purpose of supporting the health care of individual patients, new analyses to identify community wide evidence-based best practices can be enabled more easily [28].

Another important aspect is the financial impact of having access to a more complete set of information in the sense that the integration efforts may contribute for avoiding redundant tests requests and reducing time waiting for tests results to be available, both of them, with a particular impact on budgets where financial cuts are frequent [29][30]. This is particular relevant in the era of clinical governance where the ability to retrieve information quickly and accurately is considered to be very important [31].

Why is HIS integration dif cult?

Figure 2: Sometimes systems are developed without interoperability awareness coexisting in an autistic way.

However it is still a challenge to make electronic health records interoperable because good solutions to the preservation of clinical meaning across heterogeneous systems remain to be explored [32]. Consistently combining data from heterogeneous sources takes a great deal of effort because the individual feeder systems usually differ in several aspects, such as functionality, presentation, terminology, data representation and semantics [33] leading to a reality where systems coexist in a somehow autistic fashion.

While the development of health information technology, particularly electronic health records (EHR), maybe considered as a triumph for the advance of healthcare, non-interoperable clinical data systems lead to fragmented communication and incomplete records [34].

Health care is a complex environment integrating multiple participants (doctors, nurses, patients, health institutions administrations, national and regional regulation administrations and even information systems) each of them following individual agendas according to their specific tasks in the care process. Scenarios of care are diverse ranging from the most modern health institutions where facilities are in the highest of standards to improvised situations like catastrophe scenarios.

Thus, the development and implementation of EHRs faces a highly competitive environment with heterogeneous requirements from various domains involving different stakeholders [35]. Each of the stakeholder's roles and environment restrictions should be addressed attending to individual needs and characteristics, presenting information in an understandable way to each one following their different profiles of information usage.

The complexity of healthcare delivery, increases the need to have effective systems that are supported in state-of-art technologies. Health information supporting systems research should include the development and investigation of appropriate trans-institutional information system architectures, of adequate methods for strategic information management, of methods for modelling and evaluating the development and investigation of comprehensive electronic patient records, providing appropriate access for health care professionals as well as for patients and, in the broad sense including home care and health monitoring facilities [36]. The communication of EHR information is complex because much of clinical meaning is derived not from individual data values themselves but from the way in which they are linked together as compound clinical concepts, grouped under headings or problems or associated with preceding healthcare events during the act of data entry or data extraction [32]. Many distinct technological solutions coexist to integrate patient data, using differing standards and data architectures, which may difficult further interoperability [12].

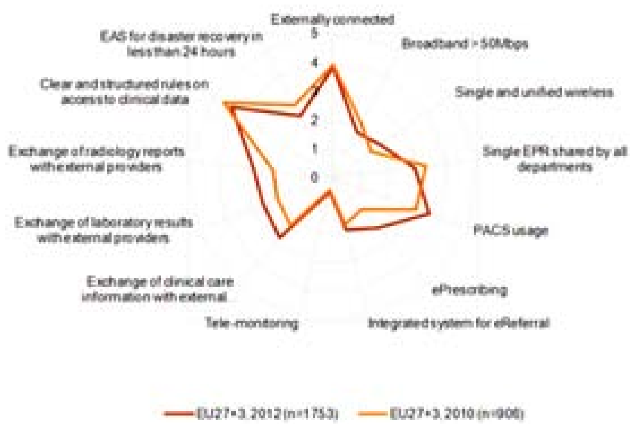


Figure 3: eHealth Profile of European acute Hospitals at EU28+2 level. 0 to 5, which respectively correspond to a 0% to 100% implementation rate.

In **??** the status of the EU average standing point regarding eHealth usage in Hospitals is described (source [37]). The exchange of data is increasing but still far from total coverage and the existence of a single EPR shared by all departments is still low, this stresses the need for additional efforts regarding the developments of additional means for making the information available where it is needed.

In general, it is pointed out that the predominant barriers to health information exchange are the need for a wider adoption of standards, security concerns, economic loss to competitors, and federated systems [34].

The EU has been promoting the adoption of interoperable EHR systems based on agreed standards throughout, first in 2004 [38] and then reenforced in 2008 with the proposal for adoption of interoperable EHR systems to support cross-border exchange of health data, in order to enhance the quality and safety' of patient care in the European space [39]. In this context, the epSOS project [40] is a relevant example where a cross-border patient summary data exchange system is being developed.

Notwithstanding this recent substantial progress, barriers continue to exist and need to be addressed in order for patients to benefit from a fully mature and interoperable eHealth systems [41].

Several standards have been proposed, to guide the development of EHR systems [32], targeting different aspects of EHR systems implementation, from high-level functional requirements, domain knowledge capturing, information structures, up to the level of value sets for field coding [42].

The major initiatives for promoting interoperable standards are now addressing issues of EHR design and implementation based in a Dual Model Architecture where clear separation between information and knowledge is enforced. Information is structured through a Reference Model containing the basic building blocks for representing any type of information of the EHR. Knowledge is represented through archetypes, which are formal definitions of clinical concepts, such as discharge report, glucose measurement or family history, in the form of structured and constrained combinations of the Reference Model building blocks providing a semantic meaning to a Reference Model structure.

This two-level modelling approach for capturing the health domain knowledge provides more flexibility by separating the informational and the knowledge concepts [43]. The interaction of the Reference Model (to store data) and the Archetype Model (to semantically describe those data structures) fosters the evolution capability of health information systems.

It is believed that semantic interoperability in healthcare can only be achieved through standardisation of data models, clinical data estrutures and terminologies [44]. The most relevant initiatives are described in the next paragraphs.

HL7 v2 messages are maybe the most widely used approach to exchange information between HIS. However these HL7 version aim mainly to support organisation and service administration along with billing purposes with few messages developed to support the shared care process itself. HL7 v2 has also revealed problems of inconsistent implementation and unsystematic growth on message segment definitions which hinder the realisation of interoperability [32]. The development of HL7 v3 face foward and aims to provide an architectural approach for semantic interoperability [45] by proposing a way of specifying the information content of messages through an information model that clarifies the definitions and ensures that they are used consistently. This is achieved by adopting a model-based specification of messages on the basis of a Reference Information Model (RIM) and a Clinical Document Architecture (CDA) that uses the concepts defined in the Reference Model and provides semantics to be used in document structures representation and hierarchies with different CDA levels allowing for different levels of granularity of presented clinical information [45].

Originating from the Synapses project [46] and evolved with inputs from the openEHR initiative [44] the **CEN/ISO 13606** is a European norm from the European Committee for Standardization (CEN) also approved as an international ISO standard designed to provide semantic interoperability in the electronic health record. The overall goal of the CEN/ISO 13606 standard is to define a rigorous and stable information architecture for communicating part or all of the electronic health record (EHR) of a single subject of care (patient) between EHR systems, or between EHR systems and a centralized EHR data repository. It may also be used for EHR communication between an EHR system or repository and clinical applications or middleware components (such as decision support components) that need to access or provide EHR data, or as the representation of EHR data within a distributed (federated) record system.

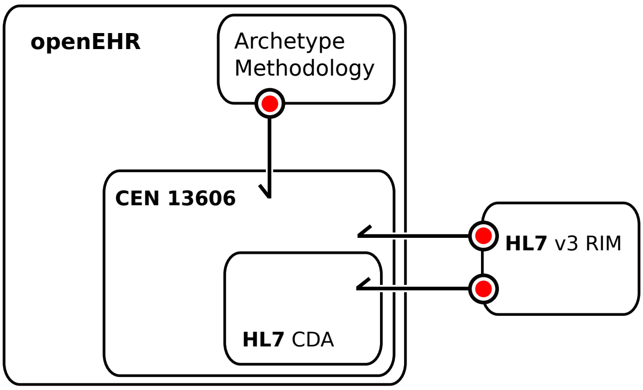


Figure 4: Schematic relationship between openEHR, CEN 13606, and HL7 CDA.

openEHR is an open standard developed with the intent of describing the management and storage, retrieval and exchange of health data in electronic health records (EHRs). It includes a Reference Model, an Archetype Model and an Service Model. These models help the creation of openEHR instances (EHR Extracts) for exchanging information. openEHR also provides the means to write semantic queries which enable querying of data independent of the underlying data model (Archetype Query Language – AQL). openEHR is becoming regarded internationally as the most complete and best-validated EHR information architecture specification [32]. The works at openEHR have had a significant influence at current standardisation efforts, both at CEN(EN-13606,EHRCom) and HL7, and future systems using any of these approaches are expected to be (to some extent) interoperable [44]. An overview of the relationship of these specifications adapted from [47] is represented in Figure 4.

Another important initiative, is the **IHE** , an industry oriented organisation whose efforts aims at creating integration profiles for particular areas covered by health information systems (e.g. Radiology, Cardiology, Laboratory). For each area it is defined an integration profile framework that promote the use of established standards such as DICOM and HL7.

Additionally, in order to successfully achieve information exchange a common terminology and coding system need to be adopted. Several coding systems co-exist in the healthcare domain but maybe the most relevant are International

Statistical Classification of Diseases and Related Health Problems (ICD) and Systematized Nomenclature of Medicine (SNOMED CT) [48]. ICD [49] was created by WHO, and is a numerical codification of diseases, signs, symptoms, procedures, complaints and external cases of injury or disease. SNOMED is a numerical codification applied in clinical records, pathology, laboratory, and decision-support systems. It comprises a standardised vocabulary system that creates a common clinical language for medical databases, containing more than 144 000 terms available in at least 12 languages.

Health information standards adoption play an important role when integration and interoperability efforts are at stake this notion is clearly stated in the eHealth report from the EC: " the use of European and international standards is a way to ensuring the interoperability of ICT solutions in general" [41].

Despite the big efforts taken through the years healthcare organisations still face the big challenge of giving all healthcare professionals complete, transparent and real time access to patient information being particularly relevant in an age where people travel more and change location more often reinforcing the notion of a great demand to create efficient integrated electronic patient records that facilitate the communication process between health professionals and the development of solutions that address these issues in the current status quo.

Achieving semantic interoperability can be considered the holly grail of EHR development. Standards usage may ease this quest but on its one they are not the silver bullet of interoperability as some are conflicting and the approaches taken to their implementation did not comply totally with the definition either because implementation changes or not enough specification is provided [42]. The articulation and inter-influences of CEN/ISO, HL7 and openEHR may ease some of this problems and pave a bit further the road to semantic interoperability a bit further.

2.2 AGENTS AND MULTIAGENT SYSTEMS

Agent theory and the development of Multi-Agent Systems (MAS) stems from the study of distributed artificial intelligence (DAI) in the 1970's. Several definitions for agents have been proposed but the one that is more widely adopted is the one from Wolldridge: "An Agent is a computer system that is capable of independent action on behalf of its user or owner".

In many cases several agents will coexist and interact in the same environment leading to a Multi-Agent System. In general, agents will be acting on behalf of users with different goals and motivations and to successfully interact, they will require the ability to cooperate, coordinate, and negotiate with each other [50].

From the general definition of Agents a series of properties were defined in order to formalize and classify the reach of agency. Agents usually enjoy the following properties [50]:

AUTONOMY: they operate without the direct intervention of humans or others, and have some kind of control over their actions and internal state.

REACTIVITY: agents perceive their environment, (which may be the physical world, a user via a graphical user interface or a collection of other agents), and respond to changes that occur in it.

SOCIAL ABILITY: agents interact with other agents (and possibly humans) via some kind of agent-communication language.

PRO-ACTIVENESS: agents do not simply act in response to their environment, they are able to exhibit goal-directed behaviour by taking the initiative.

INTELLIGENCE: agents may have reasoning, planning and learning capabilities.

MOBILITY: if they are able to move across computers in a network.

What are the main properties of Agent Systems?

When using the agent paradigm each agent can represent an individual stakeholder and they may collaborate using mutual goals. In general by coding and embodying stakeholders characteristics agents can build its basic knowledge base, which along with a representation of the environment and a set of predefine or derived plans enable the agent to pursuit the stakeholders desires. The agent based approach offers substantial advantages over current existing approaches (e.g web-services), because agents can maintain a set of workflow paths and management rules for each stakeholder within a self-contained autonomous module offering greater flexibility and dynamics when dealing with a changing environment.

2.2.1 *Agent Frameworks*

Usually Agents are created and evolve within an agent platform. An agent platform provides the necessary infrastructure for agent existence, communication and service advertisement. It can be deployed throughout several technical infrastructures from servers to mobile devices.

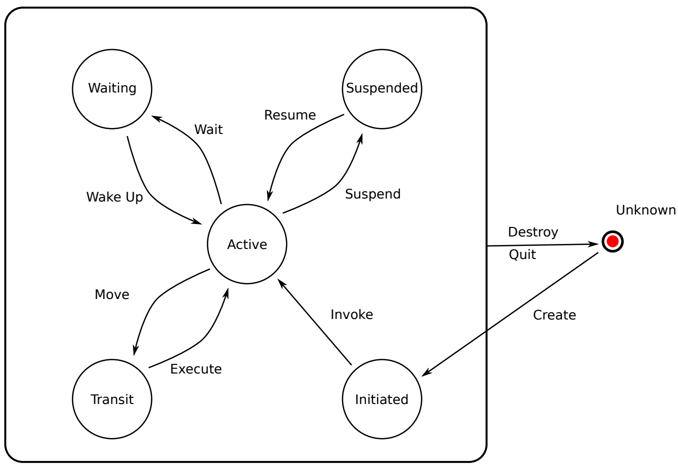


Figure 5: IEEE Fipa Agent Life Cycle.

The agent frameworks development is promoted by the IEEE Foundation for Intelligent Physical Agents (IEEE-FIPA) [51], an organisation focused on the specification for management and communication of intelligent agents. The specifications standardised by IEEE-FIPA define the basic components of an agent platform, an agent identification representation, a communication infrastructure, and agent management services(e.g. Figure 5,Figure 6).

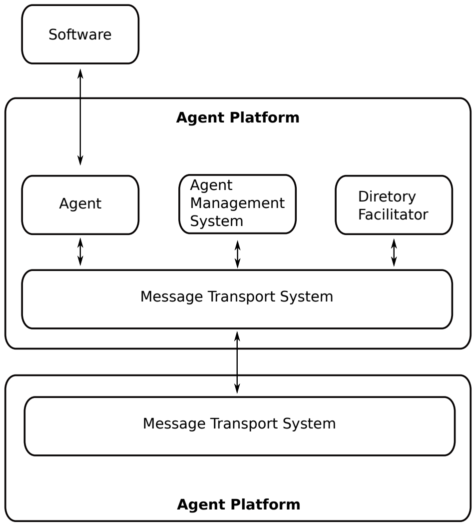


Figure 6: IEEE Fipa Platform Specification.

2.3 ADEQUACY OF APPLYING AGENT TECHNOLOGY TO HEALTHCARE PROBLEMS

In the last years MAS have emerged as an promising approach to address the issues in organising large scale and complex software systems [52]. This is also true in healthcare specially because of the particular characteristics and requisites that surround the context of healthcare software development which usually need specially tailored, flexible and adaptive solutions:

COMPLEX DATA: Many types of data and complex data structures are needed to describe the huge amounts of data resulting from observations, actions and events that take place during the care process. These include diverse items such as temperature measurements, diagnosis reports, image exams or genomic data.

DISTRIBUTED INFORMATION AND KNOWLEDGE: Information is naturally multifaceted, usually scattered around several institutions and even several systems within a single institution.

INTRICATE WORKFLOW WITH MANY PARTICIPANTS: The process of providing care is continuous and most of the time involve the articulation of multiple health professionals, from multiple departments and institutions.

Nealon and Moreno [53], suggests that Multi-agent systems may address these issues in the sense that:

How can agents be applied to Healthcare ?

- Agents can be used for information retrieval, using multiple sources, inspecting the obtained data and selecting the relevant information for a particular user (e.g. getting the most recent lab results).

- The internal state and decision process of agents can be modelled in an intuitive manner following the notion of mental attitudes. Goal-orientation means that, instead of directly requesting the agents to perform certain actions or tasks, the developer can define more abstract goals for the agents, thereby providing a certain degree of flexibility on how to achieve the goals (e.g. scheduling of an medical appointment amongst a set of institutions or health professionals.
- They are able to perform tasks that may be beneficial for the user, even if he/she has not explicitly demanded those tasks to be executed. Using this property they may find relevant information and show it to the user before he/she has to request it (e.g. alert about medication time or suggesting some reading about a patient condition).
- The components of a multi-agent system may be running in different machines, located in many different places. Each of the agents may keep part of the knowledge required to solve the problem, offering a natural way of attacking inherently distributed problems (e.g. optimising organ transplants processes among institutions).
- They are able to communicate between themselves, using some kind of agent communication language, in order to exchange any kind of information. In that way they can engage in complex dialogues, in which they can negotiate, coordinate their actions and collaborate in the solution of a problem (e.g. optimisation of patient scheduling among several institutions.

Additionally, Isern [54] has reviewed several research initiatives that further exemplified the broad spectrum of health related issues addressed by an agent based approach aggregated according to main group of problems addressed (medical data management, decision support; planning and resource allocation, remote care). Isern further identify several positive aspects derived from the use of agent in Healthcare:

- modularity
- efficency
- decentralization
- flexibility
- personalisation
- distributed planning
- monitoring
- pro-activity and security awareness

Access to health data is very important with many advantages as seen on Section 2.1.2, in most cases its completeness degree can be a critical point to make a solution more easily adopted.

In the particular aspect of integration several approaches have been developed through the years aiming to address this problem. In some cases The following are some additional examples mainly focused on information integration efforts developed aiming at providing better and more complete clinical information.

The HL7 standard is the more common standard used for addressing system heterogeneity and providing wrappers for data exchange.

Kim et al propose an agent based intelligent clinical information system for persistent lifelong electronic medical record. Functional entities are divided according to tasks and implemented as collaborating agents. These agents reside on a multiagent platform which provides communication and invocation of execution functionality. They assume the usage throughout of HL7 standard for messaging in two scenarios: patient access and inter-institutions data exchange [55].

AIDA (Agency for Integration, Archive and Diffusion of Medical Information) project aims at integrating, diffusing and archiving large sets of information from heterogeneous sources (departments, services, units, computers, medical equipments) within a single health institution. It uses a multi-agent system that include agents that are in charge of tasks such as communicating with the heterogeneous systems through XML messaging, sending and receiving information (e.g., medical or clinical reports, images, collections of data, prescriptions), managing and saving the information and answering to information requests [56].

The Medical Agent System (eMAGS) project addresses the problem of providing a flexible and scalable solution through the use of multiple cooperating mobile agents that actively access, decipher, learn and exploit the information available on various health systems. Data is exchanged through intermediary HL7 messages carried by an HL7 agent. Agent servers create or use HL7 agents to carry patient information messages around the eMAGS network. A single ontology server is used as a central repository of mappings between HL7 and subscribing database applications [57].

The MET3 system helps physicians with data collection, diagnosis formulation, treatment planning and finding supporting evidence. MET3 integrates with external hospital information systems via HL7 messages using Mirth Connect HL7 engine and runs on various computing platforms available at the point of care (e.g., tablet computers, mobile phones) [58].

In the next chapter a broad characterisation of MAS technology usage in healthcare panorama is presented regarding users addressed, deployment scenarios, implementation issues that complements the available knowledge with patterns of agent technology usage and implementation in Healthcare concluding with the identification of issues that need to be further addressed.

3 AGENTS IN HEALTHCARE - A SYSTEMATIC REVIEW

In the previous chapters we have seen that HIS development is complex and that barriers exist preventing easy systems integration both within institutions and between institutions. There were presented some examples of MAS technology being used with the purpose of overcoming some of these issues. The work that follows makes use of systematic reviews methodology to provide an overview on the MAS development in healthcare by profiling scenarios of application and the technical characteristics used for systems development.

3.1 INTRODUCTION

Due to the growth of scientific production, the role of literature reviews has been proportionally growing larger and their importance grows as a direct function of the number of documents on a topic [59] [60].

Systematic reviews represent a complementary approach to traditional literature reviews. The concept of systematic results from the methodology that defines a formal and reproducible way of synthesising the available knowledge about a particular topic. It aims to integrate empirical research in order to create generalisations [61].

By applying systematic reviews technics for identifying, indexing and summarising studies describing Agent and MultiAgent systems applied to the healthcare domain, this work aims to add an updated quantitative overview on how agent technology main characteristics are being used in the healthcare arena by gathering evidence of the broad spectre of application and deriving trends of its usage through the years.

New directions are pointed out by deriving suggestions from trends in health informatics research that can be incorporated in health related agent systems thus providing stronger implementations and better adoption of agent technology in the healthcare arena.

3.2 METHODS

3.2.1 *Search Methods And Study Selection*

Studies were searched and updated in December 2012 using the bibliographic databases Medline (via Pubmed), ISI (ISI Web of Science) and IEEE (IEEE Xplore).

The bibliographic databases were searched trough their online search interface using the queries described in Table 1 and only journal papers describing agent technology applied to healthcare were selected.

The study selection included two phases: Title and Abstract review and Full paper review and classification. The flowchart depicted in Figure 7 provides an overall view of the process.

Table 1: Search queries used in Pubmed, ISI Web of Science and IEEE Xplore.

SEARCH ENGINE	QUERY
SI WOS	(multiagent* OR multi-agent* OR agent*) AND (health* OR patient*)[1]
IEEE Xplore	(multiagent* OR multi-agent* OR agent*) AND (health* OR patient*)[2]
PUBMED	(multiagent OR multi-agent OR multiagents OR multi-agents OR agent OR agents) AND (health or healthcare OR patient) AND (Artificial Intelligence [MeSH Terms])

[1] Searched in Topic refined by Subjects: Medical Informatics; Computer Science; Artificial Intelligence; Computer Science Interdisciplinary Applications; Computer Science Information Systems.

[2] Searched in Metadata.

On the first phase Title and Abstract were read and classified according to the inclusion criteria: mentioning the use of agents and mentioning a healthcare context.

When studies' Title and Abstract content was ambiguous regarding the criteria it was nevertheless included for the second phase. After duplicate removal and application of inclusion criteria a total of 227 of the 1044 papers were selected for the next phase.

The second phase of the study selection included full paper review. The inclusion criterion in this phase was that papers should describe "Agent based systems applied to the healthcare field." Works using the word agent merely has a designation for some generic entity were not included.

A total of 102 out of the 227 papers fulfilled the inclusion criteria and were classified according to a predefined set of variables.

The papers were grouped into 79 projects to avoid distortion created by multiple papers from the same project describing the same issues. Projects id's, names (when existing) and references to the studies are presented in Table 13. Throughout the text groups of projects will be identified by a number with the preceding letter g, the projects included in these groups are described in Table 12.

For the review process an online inquiry engine MEDQUEST [63] was adapted in order to reflect the requirements of each review stage. This engine was used for each phase of the evaluation regarding data collection and pre-processing of variables.

3.2.2 *Variables De nition*

Variables examined in this review focus on a general characterisation of the use of agent technology in care provision according to main actors, scenarios and development details.

Table 2 describes the main variables observed regarding study characterisation, problems addressed, actors involved, the institutional environment, the rest of the variables regarding development are presented in line with text or

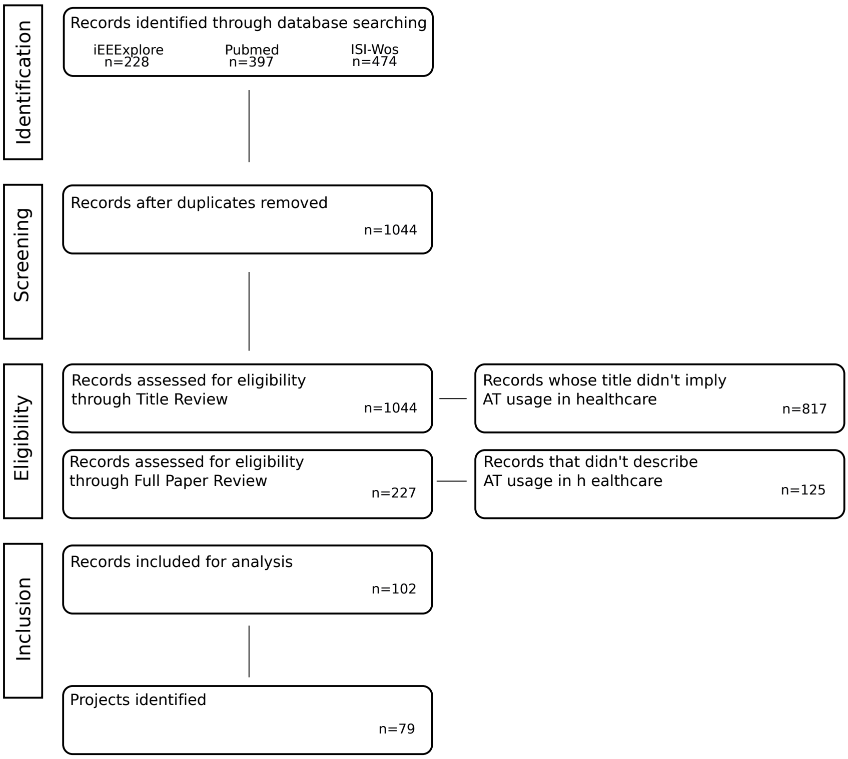


Figure 7: Workflow for studies selection (representation based on [62]).

through tables in the respective subsections. In this review values for Type of Problems addressed by projects were adapted from Isern's review [54].

3.2.3 *Time Period Considered*

To analyse time trends, we divided the total period up into five intervals. The first interval includes papers published from 2010 to 2012, the second interval from 2007 to 2009, the third from 2004 to 2006, the fourth from 2001 to 2003 and the fifth from 2000 and below. Regarding projects date classification, it was considered as project date, the date of the latest paper published.

3.3 RESULTS

3.3.1 *Studies Characterisation*

In Table 3 is represented the distribution of papers and projects along each of the time intervals considered and where it is observed increasing activity since 2004. From grouping papers into projects it is revealed that most of the projects identified (85%) have just one publication.

Table 2: Variables collected in the study.

VARIABLE	DESCRIPTION
Study	
Year	Publication year.
Country	Country of first author.
Project Name	Project name when exists or name of first author.
Type of Study	Characterisation regarding outcomes.
Problems	Characterisation of problems addressed.
Environment	
Range	What type of institutional reach is addressed.
Institutions	What are the health institutions type involved.
Actors	
End users	The main users of the system.
Special Patients	Wether a particular group of patients is addressed.
Agents Modelling	
Methodology	Wether a methodology for modelling is used.
Reasoning	Wether any kind of reasoning process is used.
Ontology	Wether ontologies are used.
Mobility	Wether mobility is used.
Agents Development	
Security	Wether security awareness is present.
Healthcare Standards	Wether health information standards are used.
Other technologies	What other technologies are used in complement.
Agents Deployment	
Number of instances	Wether single or multiple platforms are used.
Instances heterogeneity	Wether multiple types of platforms are used.
Platform used	Which platform is used.

The top five country list for Papers and Projects is shown in Table 4. Country was selected according to affiliation of first author. Spain, USA and Italy represent almost 50% of the works reviewed.

Table 3: Number and percentage of Papers and Projects along the years.

	2012/10	2009/07	2006/04	2003/00	<2000	TOTAL
Papers n (%)	31 (30)	15 (15)	28 (28)	12 (12)	16 (16)	102
Projects n (%)	27 (34)	13 (17)	20 (25)	10 (13)	9 (11)	79

The journals where the selected papers were published were grouped and classified around two areas according to their nature: Computer Science and Technology applied to Medicine. The list of journals with more than two papers is represented in In Table 5.

In Figure 8 is represented the evolution of publications according to the classification of journals as Computer Science only and Medicine and Computer

Science along the years. There is a growth in the 2000-2006 period in the computer science area reaching the hype moment around 2006 followed by growing trend in the area of applied technology in Medicine suggesting the occurrence of technology transfer to a more applied arena audience.

Table 4: Top five country list.

PAPERS		PROJECTS	
	n (%)		n (%)
Spain	19 (19)	Spain	13 (17)
USA	17 (15)	USA	13 (17)
Italy	12 (12)	Italy	8 (10)
Taiwan	8 (8)	UK	7 (9)
UK	8 (7)	Australia	4 (5)

Table 5: Journals with more than two papers.

JOURNALS	N
Journal of Medical Systems	12
IEEE Intelligent Systems	11
Lecture Notes in Artificial Intelligence	9
Artificial Intelligence in Medicine	7
IEEE Transactions on Information Technology in Biomedicine	7
Ai Communications	6
Journal of the American Medical Informatics Association	4
Lecture Notes in Computer Science	4
Expert Systems with Applications	3
Methods of Information in Medicine	3
Applied Intelligence	2
Applied Soft Computing	2
IEEE Transactions on Computational Biology and Bioinformatics	2
Computer Methods and Programs in Biomedicine	2
Computers in Biology and Medicine	2
International Journal of Medical Informatics	2

Each project was characterised according to the variables defined and the results are summarised in Table 6 and described in the next sections.

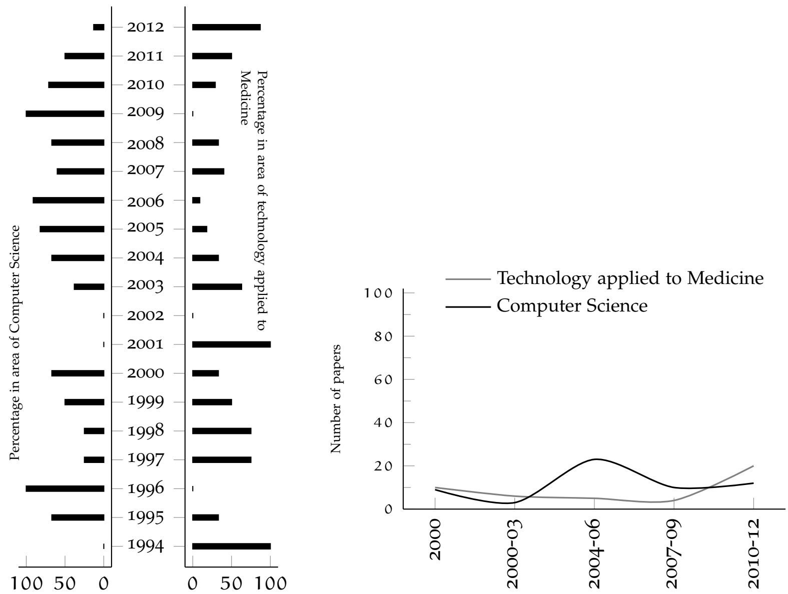


Figure 8: Distribution of papers between Computer Science area and Technology applied to Medicine area.

TYPES OF WORK: Regarding the type of work considered, all projects describes models, 66% stated to have implemented prototypes for testing purposes and only 6% stated to have deployed a system or tested it in real life settings.

PROBLEMS ADDRESSED: Regarding the problems addressed, 47% of the projects address Information Management issues followed by 41% that Decision support, 27% Coordination and Planning , 22% Remote Care and Monitoring and finally 15% Simulation/Bioinformatics. the number of projects addressing Remote Care shows a growth trend along the years. Projects where Simulation and Bioinformatics problems are addressed only appear on the more recent periods.

3.3.2 *Actors*

TYPE OF PARTICIPANTS: Regarding the entities considered for the study, Medical Doctors are present in 76% projects maintaining a steady trend along the years. Nursing Staff (30%) and the Patient (27%) follows on the list. Decision makers are considered in 19% of the projects. In 50% of the projects more than one type of actor is taken in consideration.

PARTICULAR GROUP OF PATIENTS: From the reviewed projects, 34% of them are aimed at particular group of patients or diseases specificities (*go*), and have the following distribution: Elderly (5%); Alzheimer Patients (3%); Brain Tumor Patients (3%); Generic Chronic Diseases (3%); Depression (3%); Emergency patients (3%); Epilepsy (3%); Organ Donors/Recipients (3%); Acute Myeloid Leukaemia (1%); Cardiac Surgery patients (1%); Deep Venous Thrombosis (1%); Influenza Infection (1%); Intensive Care (1%); Lower Urinary Tract Dysfunction Patients (1%); Neuronal Dysfunctions (1%); Malaria Patients (1%).

Table 6: Frequencies and percentages for each variable analysed among the projects reviewed.

VARIABLES	n	%	2012/10 n=27	2009/07 n=17	2006/04 n=21	2003/00 n=10	<2000 n=9	PROJECTS
Outcomes								
Model	79	(100)	27 (100)	13 (100)	20 (100)	10 (100)	9 (100)	g1
Prototype	52	(66)	18 (67)	9 (69)	12 (60)	6 (60)	7 (78)	g2
Deployed	5	(6)	2 (7)	0 (0)	3 (15)	0 (0)	0 (0)	g3
Problems								
Decision Support	32	(41)	9 (33)	8 (62)	8 (40)	3 (30)	4 (44)	g4
Information Management	37	(47)	12 (44)	3 (23)	10 (50)	6 (60)	6 (67)	g5
Remote Care	17	(22)	9 (33)	5 (39)	1 (5)	1 (10)	1 (11)	g6
Coordination	21	(27)	6 (22)	2 (15)	6 (30)	4 (40)	3 (33)	g7
Simulation Bioinformatics	12	(15)	7 (26)	3 (23)	2 (10)	0 (0)	0 (0)	g8
Actors								
Doctors	60	(76)	21 (78)	9 (70)	15 (75)	8 (80)	7 (78)	g9
Nurses	24	(30)	12 (44)	2 (15)	0 (0)	5 (50)	5 (56)	g10
Patient	21	(27)	10 (37)	1 (8)	5 (25)	3 (30)	2 (22)	g11
Decision Makers	15	(19)	5 (19)	4 (31)	4 (20)	1 (10)	1 (11)	g12
Range								
Institutional	37	(47)	14 (52)	3 (23)	7 (35)	7 (70)	6 (67)	g13
Regional National	40	(51)	13 (48)	9 (69)	13 (65)	2 (20)	3 (33)	g14
International	4	(5)	0 (0)	1 (8)	3 (15)	0 (0)	0 (0)	g15
Scenarios								
Hospital	46	(58)	18 (67)	6 (46)	7 (35)	9 (90)	6 (67)	g16
Primary Care	14	(18)	6 (22)	2 (15)	3 (15)	0 (0)	3 (33)	g17
Laboratories	8	(10)	4 (15)	1 (8)	0 (0)	1 (10)	2 (22)	g18
Home	13	(17)	7 (26)	2 (15)	2 (10)	1 (10)	1 (11)	g19
Generic Health Institution	22	(28)	7 (26)	4 (30)	8 (40)	1 (10)	2 (22)	g20

3.3.3 *Environment*

AREA COVERED: Regional/National area coverage is the most common among projects (51%), followed by Institutional coverage (47%). International coverage is residual (5%).

TYPE OF INSTITUTIONS INVOLVED: In most of the projects (58%) Hospital is the main healthcare institution addressed, a great number refer generic health institutions (28%) followed by Primary Care (18%). Regarding proximity care, some projects (17%) include Home in the systems scenario. 48% of the Projects include more than one institution.

3.3.4 *Cross-references of variables*

Tables 7, 8, 9 show the cross-referencing between Types of Problems with Main Actors and Scenarios of Application and between Actors and Main Health Institutions with the purpose of identifying the most common profiles.

TYPES OF PROBLEMS VS RANGE: From the cross-reference of type of problems and the range of application, at the Institutional setting, Information Management (43%) and Decision support (41%) have the higher percentage of cases and at the Regional level information Management has the higher representation (63%).

Looking individually to each type of problem, Decision support, Information Management, Coordination and Remote Care/Monitoring have a higher incidence at the Regional level while Simulation and Bioinformatics is more represented at the Institutional Level.

The more frequent profile is related to Information Management issues and Regional settings (34%).

Table 7: Cross-reference of Problems vs Range.

	Institutional	Regional	International
Decision Support	15 (48) (41) (21)	17 (55) (43) (23)	2 (7) (50) (3)
Information Management	16 (43) (43) (22)	25 (68) (63) (34)	2 (5) (50) (3)
Coordination	9 (43) (24) (12)	13 (62) (33) (18)	2 (10) (50) (3)
Remote Care Monitoring	8 (50) (22) (11)	9 (56) (23) (12)	0 (0) (0) (0)
Simulation Bioinformatics	6 (75) (16) (8)	2 (25) (5) (3)	0 (0) (0) (0)

Data is shown as: number of cases (% within line)
(% within col) (% within total)

TYPES OF PROBLEMS VS ACTORS: Medical Doctors are mainly addressed in problems regarding of Information Management (52%) and Decision Support (42%) and Nurses in problems of Coordination (42%) and RemoteCare/Monitoring (46%).

Policy Makers are more involved in Decision Support (73%). The Patient is more present in Information Management (52%) and RemoteCare/Monitoring (48%) issues.

The more frequent profile is Information Management and Doctors (41%).

ACTORS AND MAIN HEALTH INSTITUTIONS: Looking to how actors relate to health institutions it is observable that all actors are more represented in Hospitals. On the particular case of Patients, Home also has a relevant participation (61%).

The more frequent profile is related to Doctors and Hospitals (79%).

Table 8: Cross-reference of Problems vs Actors.

	Doctors	Nurses	Patient	Policy Maker
Decision Support	25(78) (42)(34)	7(22) (29)(10)	6(19) (29)(8)	11(34) (73)(15)
Information Management	31 (91) (52) (43)	8 (24) (33) (11)	11 (32) (52) (15)	6 (18) (40) (8)
Coordination	14 (70) (23) (19)	10 (50) (42) (14)	9 (45) (43) (12)	5 (25) (33) (7)
Remote Care Monitoring	15 (88) (25) (21)	11 (65) (46) (15)	10 (59) (48) (14)	2 (12) (13) (3)
Simulation Bioinformatics	6 (67) (10) (8)	1 (11) (4) (1)	0 (0) (0) (0)	4 (44) (27) (6)

Data is shown as: n cases (% within line) (% within col) (% within total)

Table 9: Cross-reference of Locations vs Actors.

	Doctors	Nurses	Patient	Policy Maker
Hospital	42 (91) (91) (79)	19 (41) (91) (36)	11 (24) (61) (21)	9 (20) (100) (17)
Primary Care	12 (86) (26) (23)	7 (50) (33) (13)	7 (50) (39) (13)	4 (29) (44) (8)
Home	10 (77) (22) (19)	8 (62) (38) (15)	11 (85) (61) (21)	1 (8) (11) (2)

Data is shown as: n cases (% within line) (% within col) (% within total)

3.3.5 *Implementation Issues*

The next subsections will focus on several aspects of agent based system implementation, describing the practical aspects of the reviewed projects.

3.3.5.1 *Agent Systems Modelling*

Only 13% of the projects have reported to make use of an agent orient ed methodology for system analysis and design or a agent oriented notation language (*g21*). From these, 4% refer to use GAIA [64], the rest make use of one of the following:

Prometheus [65], MASE [66], Electronic Institutions[67]. For describing processes 4% use AUML [68] as a notation language.

From the reviewed projects, 38% describe the use of a reasoning process (*g22*) and 35% refer or describe the use of Ontologies (*g23*).

Mobile agents are present in 11% of the projects (*g24*).

3.3.5.2 *Agent Systems Development*

SECURITY: Security awareness is reduced, being only present in 28% of the projects which is a bit lower when filtering by projects that implemented prototypes (23%). A lower geographical range is related to a lower value of awareness prevalence Institutional 14% vs International (50%), Regional/National settings (45%). Security awareness also varies according to types of problems (Table 10) where 46% of the projects regarding Information Management refer some kind of security measure, on the other end is Decision Support with only 16% and Simulation/Bioinformatics (0%).

Table 10: Cross-reference of Security awareness vs Problems.

	PRESENT n (%)	NOT PRESENT n (%)
Information Management	17 (46)	20 (54)
Coordination	8 (38)	13 (62)
Remote Care And Monitoring	6 (35)	11 (65)
Decision Support	5 (16)	27 (84)
Simulation/Bioninformatics	0 (0)	12 (100)

From the ones that refer to enforce some kind of security measure the most relevant proposals aim at Integrity, Non Repudiation and Authentication by protecting the infrastructure, message content, agent code and information access. The most common approaches rely on secure communication channels, Digital signatures, Hashes, Public Key Infrastructure (PKI) and Role Based Access Control (RBAC) for managing information access levels (*g25*).

HEALTHCARE INFORMATION MANAGEMENT: HIS related standards usage has low incidence in the reviewed projects. Only 18% have referred to use one or more Healthcare related standards. HL7 is used in 14% of the projects, SNOMED CT [48] and ICD [49] are referred in 4% and UMLS [69] in 3% (*g26*). Patient access is present in 22% of the projects (*g27*).

3.3.5.3 *Agent Systems Deployment*

Regarding agent organisation the majority of systems has more than one agent (89%) and for those with prototypes it is used mainly a single instance of the platform (54%).

Regarding Agent Management Systems (AMS), in total, the self-developed frameworks are more frequent (44%) followed by the Tilab Jade system which is by far the most commonly used agent framework (33%) of the projects. However It is interesting to notice the reduction in self developed frameworks along the years contrasting with an increase on Jade usage. The projects reporting deployed systems use Jade framework. In Table 11 is shown the agent frameworks identified and their usage evolution along the years.

Regarding the use of other technologies, 8% of the projects refer to integrate or make use of Webservices as complementary technology (*g28*).

Table 11: Agent frameworks in projects with prototypes.

PLATFORM			2012/10	2009/07	2006/04	2003/00	<2000	PROJECTS
	n	%	n=18	n=9	n=12	n=6	n=7	
Selfdeveloped	23	(44)	4 (22)	4 (44)	6 (50)	4 (67)	5 (71)	g29
Jade [70]	17	(33)	11 (61)	2 (22)	3 (25)	1 (17)	0 (0)	g30
Jack [71]	2	(4)	0 (0)	2 (22)	0 (0)	0 (0)	0 (0)	g31
Aglets [72]	2	(4)	0 (0)	0 (0)	0 (0)	0 (0)	2 (29)	
Salsa[73]	1	(2)	1 (6)	0 (0)	0 (0)	0 (0)	0 (0)	
Jude [74]	1	(2)	0 (0)	0 (0)	1 (8)	0 (0)	0 (0)	
Madkit [75]	1	(2)	0 (0)	1 (11)	0 (0)	0 (0)	0 (0)	
Soar [76]	1	(2)	0 (0)	0 (0)	1 (8)	0 (0)	0 (0)	
FIPA SMART [77]	1	(2)	0 (0)	0 (0)	0 (0)	1 (17)	0 (0)	
Brahms, Agent iSolutions [78]	1	(2)	1 (6)	0 (0)	0 (0)	0 (0)	0 (0)	
Open Agent Architecture [79]	1	(2)	0 (0)	0 (0)	1 (8)	0 (0)	0 (0)	
Repast Symphony suite [80]	1	(2)	1 (6)	0 (0)	0 (0)	0 (0)	0 (0)	

3.4 DISCUSSION

Although we feel that grouping studies into projects is essential to decrease the bias of multiple publications of the same project, on some of the studies it was difficult to determine if they were describing the same project or not. Also the decision to include only journals may lead to leaving out works only published in conference proceedings, however given that usually journal papers represent more mature and complete description of research this criteria was used for studies triage.

Gathered evidence reveals a growing tendency on publications regarding the application of AT in the Healthcare scenario with a slightly decrease around 2007/09. This follows the notion of the widespread adoption of agent technology also evidenced by a growth along the years in most of the set of problems classified. Areas of publication also show the evolution of the technology first with a higher manifestation on the Computer Science area and later an increasing tendency for publication in journals regarding applied technology to the healthcare field.

An example of the dynamics is Remote Care and Monitoring that became an area of increased activity in the last years reflecting the focus given nowadays to the ageing population and to the extension of care provisioning outside standard locations. This is also reflected by the increasing presence of Home in the scenario variable.

Problems addressed are heterogeneous however decision support and information availability are the most common issues addressed by agents, this areas are highly interrelated as for decision support information is vital. This is evidenced by cross-reference problems and geographical range where there is high percentage of decision support projects in single institutions while information availability issues are higher in scenarios of multiple institutions. While in the first case information may be easily accessed within the institution in wider scenario difficulties may rise for information access and consequently create obstacles for decision support applications.

Regarding the actors in the Healthcare a high percentage of projects involving Doctors is expected nonetheless it is worth to notice that other players (nurses, patients and decision makers) are becoming more present in recent years. Additionally the fact that a high number of projects address particular diseases or conditions shows that AT can embody a fine grain regarding actor's specificity.

Being healthcare naturally involving multiple interveners the higher frequency of Multi agent systems vs Single agent systems is natural.

Jade is the widely used AMS, this is particularly evident in the last years and may result from its maturity, active enhancements and support available. It is important to notice that there is a decreasing percentage of the number of self developed AMS through the years showing the increasing maturity of the AMS available.

Agent oriented methodologies are still rarely used, an approach that might contributed for a wider adoption could be the development of a more direct articulation with agent deployment frameworks enabling a more quick transition from modelling to implementation.

HIS related standards should be considered more seriously when dealing with healthcare information specially when there is the need for exchanging data. A wider adoption could provide stronger grounds for deployment in real scenarios.

Although present in some systems, security awareness or implementation should be stronger, this would also allow for easier adoption of the technology in the healthcare domain.

Deployed applications experiences and impact in the daily practice evaluation are not very common. A few deployed projects in real scenarios are reported and most of them are related to Home monitoring. Efforts should be made on order to improve agent technology visibility in the enterprise community. Additionally, it is our belief that a more systematic and standard description of agent systems clearly identifying the health actors, environment specificities and implementation details would enable a deeper comparison between systems and provide the outside community with a sharper understanding of the advantages of applying AT in healthcare.

3.5 OUTCOMES

- **P. Vieira-Marques**, R. J. Cruz-Correia, and Sergi Robles, Profiling Roles and Scenarios of Agent application in Healthcare - a Systematic Review.VIII Workshop on Agents Applied in Health Care, held in conjunction with the 14th Conference on Artificial Intelligence in Medicine (AIME 2013)[9].

Table 12: Projects Groups.

GROUPS	PROJECTS
g0	6, 10, 18, 19, 24, 25, 26, 28, 30, 35, 36, 37, 38, 39, 43, 45, 46, 51, 52, 55, 56, 58, 68, 71, 72, 75
g1	All
g2	1, 3, 4, 5, 6, 9, 10, 12, 14, 15, 17, 19, 20, 21, 22, 23, 24, 25, 26, 29, 30, 32, 33, 35, 37, 38, 39, 41, 43, 44, 50, 51, 52, 53, 54, 55, 57, 58, 60, 61, 62, 63, 66, 67, 69, 71, 72, 73, 74, 75, 77, 79
g3	10, 26, 41, 52, 58
g4	1, 3, 4, 9, 19, 20, 21, 22, 23, 24, 29, 30, 33, 37, 38, 39, 40, 43, 47, 51, 53, 54, 56, 57, 60, 61, 63, 65, 71, 73, 75, 79
g5	1, 2, 3, 4, 7, 9, 11, 13, 14, 16, 22, 24, 29, 31, 34, 41, 42, 44, 45, 49, 50, 56, 57, 58, 59, 61, 63, 64, 66, 67, 68, 71, 73, 74, 76, 77, 78
g6	6, 11, 12, 14, 15, 16, 19, 25, 26, 28, 35, 36, 38, 40, 52, 62, 72
g7	2, 9, 10, 14, 16, 19, 21, 33, 43, 45, 46, 48, 49, 50, 63, 67, 69, 70, 71, 76, 79
g8	5, 8, 17, 18, 20, 23, 27, 29, 32, 37, 55, 60
g9	1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 22, 24, 25, 26, 30, 31, 34, 35, 36, 37, 38, 40, 41, 42, 43, 44, 45, 47, 50, 51, 52, 53, 54, 55, 56, 57, 60, 61, 62, 64, 65, 66, 67, 68, 69, 71, 73, 74, 75, 76, 77, 79
g10	6, 9, 10, 12, 13, 14, 15, 16, 17, 19, 22, 25, 26, 35, 62, 65, 66, 69, 70, 71, 72, 75, 76, 79
g11	3, 9, 13, 14, 15, 16, 19, 21, 25, 26, 28, 48, 50, 52, 56, 58, 62, 63, 64, 72, 76
g12	16, 18, 20, 22, 23, 29, 33, 39, 40, 42, 43, 46, 57, 65, 71
g13	1, 4, 6, 8, 9, 10, 11, 12, 14, 17, 18, 20, 26, 28, 36, 37, 38, 41, 44, 51, 53, 54, 58, 60, 61, 65, 66, 67, 68, 69, 70, 71, 75, 76, 77, 78, 79
g14	2, 3, 7, 9, 11, 13, 15, 16, 19, 21, 22, 23, 24, 25, 29, 30, 31, 33, 34, 35, 39, 40, 41, 42, 43, 44, 45, 46, 48, 49, 50, 52, 56, 57, 59, 63, 64, 72, 73, 74
g15	27, 71, 69, 83
g16	2, 4, 6, 7, 9, 10, 11, 12, 14, 15, 16, 18, 19, 21, 22, 23, 24, 25, 26, 30, 35, 36, 38, 40, 41, 43, 44, 45, 54, 57, 60, 61, 62, 63, 64, 65, 67, 68, 69, 70, 71, 73, 74, 75, 77, 79
g17	9, 13, 16, 19, 22, 23, 24, 40, 45, 50, 52, 72, 74, 79
g18	7, 11, 16, 19, 22, 65, 72, 79
g19	3, 11, 13, 15, 16, 19, 25, 28, 35, 52, 58, 62, 72
g20	1, 8, 9, 11, 14, 17, 20, 29, 31, 34, 39, 42, 46, 48, 49, 51, 53, 56, 59, 66, 76, 78
g21	4, 10, 14, 19, 22, 26, 29, 42, 52, 65
g22	6, 8, 9, 10, 12, 14, 18, 19, 22, 26, 28, 29, 30, 31, 33, 36, 38, 39, 40, 43, 51, 52, 53, 57, 65, 70, 72, 73, 75, 79
g23	3, 4, 8, 9, 12, 14, 16, 19, 21, 24, 29, 30, 32, 38, 39, 42, 43, 44, 48, 49, 51, 52, 57, 63, 71, 75, 76, 79
g24	3, 7, 11, 14, 15, 31, 41, 44, 74
g25	7, 11, 13, 14, 16, 19, 21, 22, 24, 25, 34, 41, 42, 45, 46, 49, 59, 62, 63, 66, 74, 77
g26	2, 4, 16, 21, 25, 29, 40, 42, 44, 61, 64, 73, 77, 78
g27	7, 13, 14, 16, 19, 21, 25, 34, 45, 48, 50, 59, 63, 64, 65, 69, 77
g28	1, 10, 16, 19, 31, 64
g29	10, 17, 23, 24, 32, 33, 35, 38, 43, 44, 53, 54, 55, 58, 61, 62, 66, 69, 71, 72, 73, 75, 79
g30	1, 3, 4, 6, 9, 12, 15, 19, 21, 22, 25, 26, 39, 41, 52, 57, 63
g31	5,14, 20, 29, 30, 37, 50, 51, 60, 67, 74, 77

Table 13: Projects identification and references.

P.1	Wu[81]	P.28	Ab Aziz[82]	P.55	Fernandez[83]
P.2	UP-ColBPIP[84]	P.29	DECIMAS[85][86]	P.56	Druzovec[87]
P.3	ONTORIS[88]	P.30	HealthAgents[89]	P.57	COSMOA[90]
P.4	MET3[58]	P.31	MOBIFLEX[91]	P.58	PSN[92]
P.5	Paoletti[93]	P.32	Pappalardo[94]	P.59	Gritzalis[95]
P.6	Nouira[96]	P.33	Lopez[97]	P.60	Bovenkamp[98]
P.7	Nikooghadam[99]	P.34	Choe[100]	P.61	MIRRORS[101]
P.8	MVP[102]	P.35	Chan[103]	P.62	TIISSAD[104]
P.9	HeCaSe2[105][106][107]	P.36	Vazquez[108]	P.63	HeCase[109]
P.10	ALZ-MAS[110][111]	P.37	Ndiaye[112]	P.64	Kim[55]
P.11	Chen[113][114][115][116]	P.38	Lee[117]	P.65	Godo[118]
P.12	Catarinucci[119]	P.39	Koum[120]	P.66	JAFDIS[121]
P.13	GUISM[122]	P.40	Ji[123]	P.67	TOMAS[124]
P.14	CHIS[125][126][127][128][129]	P.41	MAID[3][1]	P.68	UPRR[130]
P.15	MADIP[131][132]	P.42	Ulieru[133]	P.69	Marinagi[134]
P.16	LuMiR[135]	P.43	CARREL[136][137][138]	P.70	Decker[139]
P.17	Micro-Gen[140]	P.44	eMAGS[57]	P.71	PATMAN[141][142][143]
P.18	Laskowski[144]	P.45	OTMA[145]	P.72	MOBCare[146]
P.19	K4Care[147]	P.46	InCA[148]	P.73	R2DO2[149]
P.20	Dunn[150]	P.47	ASPIC[151]	P.74	Murphy[152]
P.21	De Meo[153]	P.48	CASCOM[154]	P.75	GUARDIAN[155][156][157][158]
P.22	ADRMonitor[159]	P.49	Vazquez-Salceda[160]	P.76	CASIS[161][162][163]
P.23	Xuyan[164]	P.50	Reed[165]	P.77	Rabbani[166]
P.24	HealthAgents2[167]	P.51	Letia[168]	P.78	HECS[169]
P.25	Vassis[170]	P.52	CHS[171]	P.79	AADCare[172]
P.26	MaRV[173]	P.53	Hara[174][175]		
P.27	Bonjean[176]	P.54	Guyet[177]		

Part II

THE JOURNEY FROM INTRA TO INTER INSTITUTIONAL HEALTH DATA INTEGRATION

```
Log Entry #2 : By now this journey has provided pictures
of how healthcare is knowledge driven, that information
should be available and that this is a complex task. Also
that agent technology is actively addressing relevant
health information related issues. In the following stages
the efforts taken onwards exploring further this technology
for enhancing data availability are described in detail.
```

4 AGENT BASED INTRA-INSTITUTION DATA INTEGRATION

After describing the literature and profiling the use of agents in the healthcare domain we describe in this chapter an agent based system for integration of legacy health information systems within a single health institution. The MAID system was developed with the intent of providing enhanced access to health information from several departments throughout a major Hospital.

4.1 SCENARIO

With approximately 1.350 beds, Hospital S. João (HSJ) is the reference hospital for the North of Portugal and the second largest in the country. This university hospital employs around 800 staff doctors and another 400 as trainees, from over 50 medical specialities. Similarly to all other Portuguese public hospitals, at

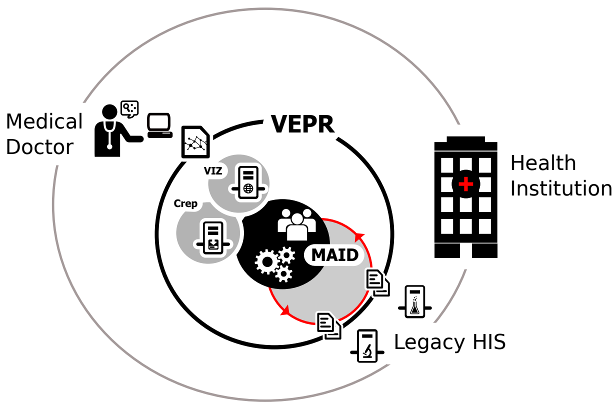


Figure 9: Generic model of MAID - Multi Agent system for the Integration of health Data.

HSJ, patient's information resources are based on SONHO, a system developed by the governmental agency IGIF (Instituto de Gestão Informática e Financeira da Saúde) that manages data related to hospital stays, consultations and emergency visits. Although this system has the capacity to manage some general clinical data (in a module called SAM – Sistema de Apoio ao Médico) it has reduced flexibility regarding the use of these data by health professionals and lacks specificity regarding the needs of different medical specialities. This may explain why SONHO is still mostly used for administrative (e-g. consultations scheduling) and financial (e.g. Diagnosis Related Groups) purposes rather than for the care of patients.

Not surprisingly, nearly two thirds of its major clinical departments have implemented or acquired at least one information system to record specific data focused on the daily management of their patients and on their own medical research interests. Unfortunately, these departmental information systems (DIS) use many different data structures, database management systems, ontologies, communication protocols, file formats for reporting clinical results and user authentication systems and are mainly available to the departmental staff. In addition, many of these systems are connected to medical devices (e.g. monitoring, imaging or lab instruments), which increase the overall complexity.

This real scenario includes unrelated applications managing relevant pieces of clinical information and duplicated data scattered over several databases. Therefore, the vast majority of patient's data could only be shared using paper records, conducing to an inefficient flow of information translated into high administrative costs both in time and staff.

Multi-agent technologies offered a natural and transparent way of mapping this information flow and corresponding human actions (e.g. locate, pull, copy, transport and store clinical reports) during the daily hospital routine. In Figure 9 it is presented a generic model of the proposed system aimed at integrating health data within a single health institution.

4.1.1 *Actors*

The actors proxied by this system are doctors that want to access information from other departments during their daily routine, systems that provide information and the staff members that participate on the process of making information available at the point of care.

Doctors may interact with the system trough a single web interface. External systems are passive players awaiting for requests and the auxiliary staff is incorporated in the system by having their actions mimicked by agents' actions.

The focus of the system regarding the actual workflow for information sharing within the hospital is depicted in Figure 10 which illustrates a common process of disseminating the information within the hospital.

In scenario A (delivery of reports by producer), a laboratory technician when finalises the lab results, asks an auxiliary to deliver the results to the requesting department. The results are received by the administrative personal who finally distributes this information to the patient doctor.

In scenario B (delivery of reports by requester), a doctor or an administrative auxiliary on his behalf, periodically checks the lab for the readiness of the results, when they are finally ready scenario A is triggered.

4.1.2 *Environment*

The system's environment include many departmental islands of information, addressing multiple medical specialties, based on multiple programming paradigms and technologies. The departmental information systems integrated at the early stages are described in Table 14. This list was further extended after deployment with additional departmental information systems: Psychiatry; Gynecologic Endoscopy; Anaesthesiology; Thoracic Surgery; Clinical Nephrology and Clinical Haematology.

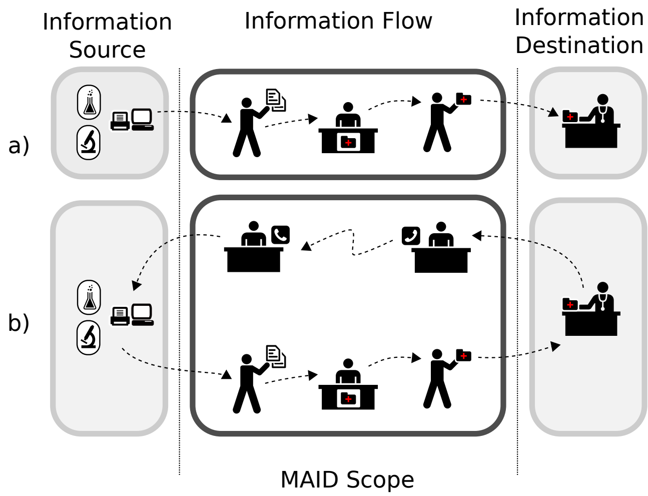


Figure 10: Scenarios of actual processes for making data accessible within the health institution: a) Delivery of reports by producer; b) Deliver of reports by requester.

The system is available on the Hospital private network and interacts both with the clinicians and the DIS. Doctors can access the system through an integrated web interface.

The system sense existing changes on the DIS data by interaction with a set of predefined interfaces that provide a wrapper hiding DIS heterogeneity.

Table 14: Integrated systems in the early stages of the project.

DIS	TECHNOLOGY USED	WEB ACCESS	TYPE OF DATA
Anatomical Pathology	Microsoft Access	No	Exams
Cardiology	Visual Basic	Yes	ECGs and discharge letters
Clinical Pathology	Visual Basic / PHP	Yes	Biochemical and microbiologic results, haematological cell counts
Gastroenterology	Delphi	No	Upper and lower endoscopy reports
Gynaecology	Java	Yes	Endoscopy reports
Immune-alergology	Delphi	No	Pulmonary function tests
Immune-haemotherapy	Visual Basic	Yes	Coagulation studies, molecular biology, immune-hematologic results
Intensive Care	Java	Yes	Discharge letters
Pneumology	Java	Yes	Bronchoscopy reports and discharge letters
Paediatric-Gastroenterology	Delphi	No	Upper and lower endoscopy reports
Obstetrics	Java	Yes	Discharge letters, birth reports, exam reports

4.1.3 *Data Integration Model*

Two levels of integration were taken in consideration for this system: data level and document level. In the first level resides the access to patient administrative data accomplished by accessing central SONHO database. In the second level resides the integration of data produced in the several departments (clinical data).

The integration with SONHO database is used to retrieve patient administrative information and also for cross-referencing the identifications provided by the departments.

For the departmental data integration it was chosen a document based approach. The document based integration with the DIS was selected over a more standard approach like HL7 wrappers to data. This option was largely forced by the significant amount of time and money that would be necessary to address the high level of heterogeneity regarding data structure and presentation formats exhibited and the consequent need to create individual systems infrastructure profiles and the mapping for stored data. Hence, the document level was adopted for clinical data integration, usually using pdf or html files, that are collected by the central system. Each query to the departments will generate a snapshot of the more recent produced reports (e.g. discharge letter or a lab report) which are associated to the department of origin in the user interface.

4.1.4 *Patient Identi cation*

One crucial aspect for systems integration is the use of a common patient identification. Given that in SONHO, several patient identifiers and contact numbers are stored (these are described in Table 15), the task of selecting one identifier that would be the best representative was a major barrier to overcome. Particularly, because their use doesn't follow a common pattern among departments and while some use the process number (*num_processo*) others use the encounter number (*num_episodio*) for patient identification. Additionally given that the national patient identification (*num_utente*) and the process number (*num_processo)* are not always present in SONHO it was decided to use the *num_sequencial* as the main patient identifier in MAID and then use the central administrative database to convert the used numbers to this.

Table 15: Types of patient identifiers used by HSJ information systems.

PATIENT IDENTIFIERS	DESCRIPTION
num_utente	identifier of Portuguese NHS (Sistema Nacional de Saúde)
num_sequencial	unique internal patient identifier of SONHO for each hospital
num_processo	clinical patient record identifier for each hospital
num_episodio + cod_modulo	number and type (outpatient, inpatient, emergency, laboratory, radiology) of encounter with the hospital

4.2 SYSTEM MODEL

The general design of MAID is described in Figure 11. It explores some of the characteristics of the multi-agent systems paradigm namely their independence, autonomy and pro-activeness. MAID main entities were designed to cooperate and undertake the necessary actions in order to build a virtual electronic patient

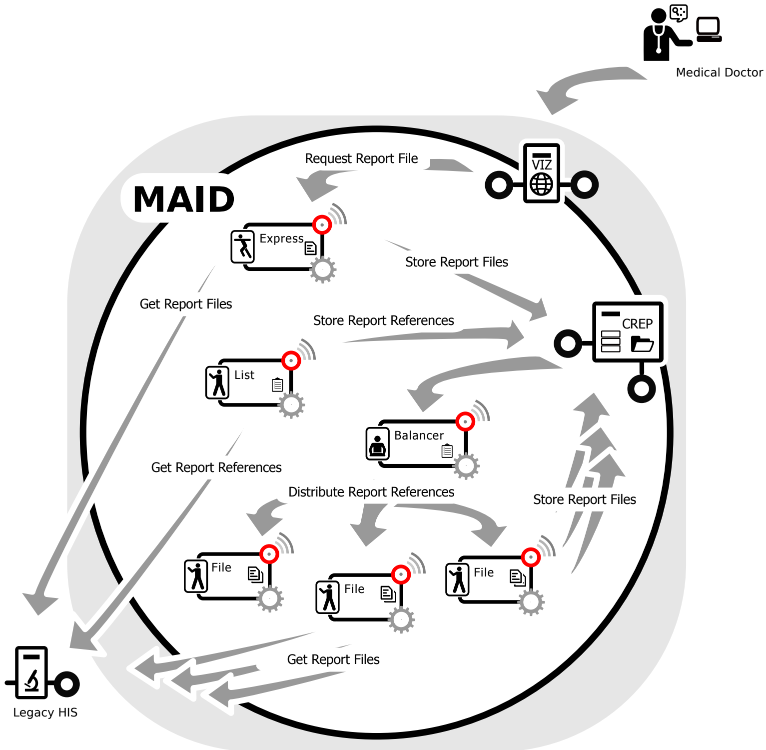


Figure 11: Agents community with the main actions performed during the process of making information available from legacy HIS to medical doctors.

record, making existing information available throughout the hospital, within a reasonable time frame, trough a single point of access to all medical doctors.

The automatic document retrieval process is split into two independent action processes performed by a set of agents: firstly, the retrieval of report references consists of the questioning of each DIS for new clinical reports and the eventual retrieval of their references (List Agents); secondly, the retrieval of report files consists on the actual retrieval of the correspondent clinical reports files from the DIS to the central repository (balancer and file agents). A particular case of this last action corresponds to the immediate retrieval of a clinical report that is not yet available centrally but has meanwhile been requested by an user.

The described processes are mapped into behaviours that define agent's singular characteristics and individual actions. The articulation of these behaviours describes the necessary steps that the agents must undertake in order to accomplish their purpose of collecting new data. All agents act on an asynchronous and autonomous way. The CRep module include the database and filesystem and VIZ represents the web interface used for making information available to the end users.

4.2.1 *Setup Process*

As the system starts, a series of actions have to be taken in order to correctly launch the system's core agents. A set of property files is used to define the startup variables. The control agent is responsible for creating all the agents in the system and to respond to external monitoring requests messages from a NAGIOS system [178] thus enabling remote monitoring of system execution and resource usage. Control agent exposes a listening service for incoming requests purposes. The execution flow is depicted in Figure 12.

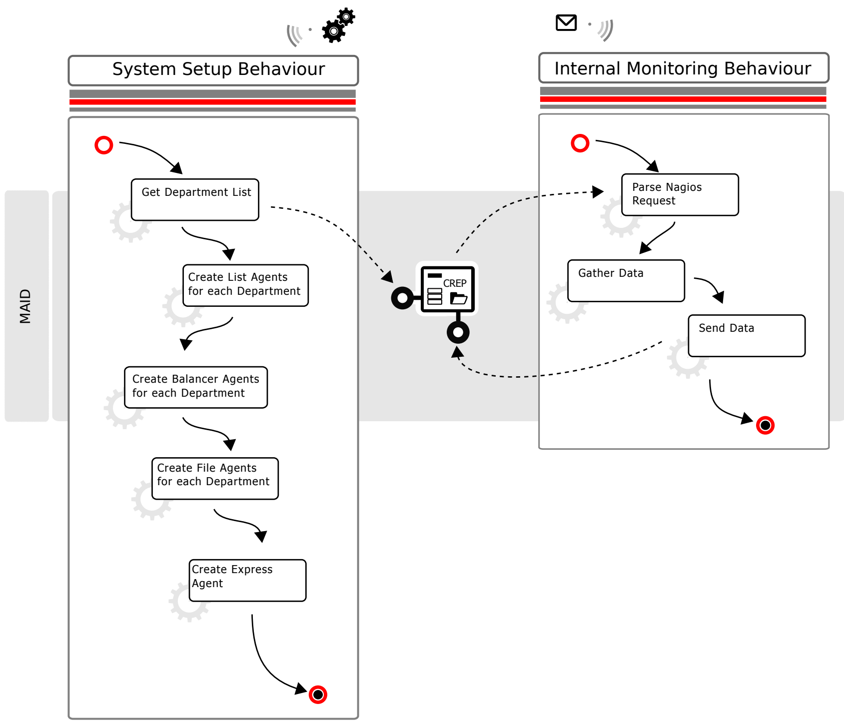


Figure 12: Control Agent main actions and interactions during the system startup.

4.2.2 *Reports Reference List Retrieval Process*

The first stage of making a report available is the retrieval of report references, this is accomplished by List Agents whose model and main characteristics are described bellow.

4.2.2.1 *List Agents Model*

An individual List Agent is assigned to each of the Departments integrated. These agents regularly survey the DIS looking for new clinical reports. For each

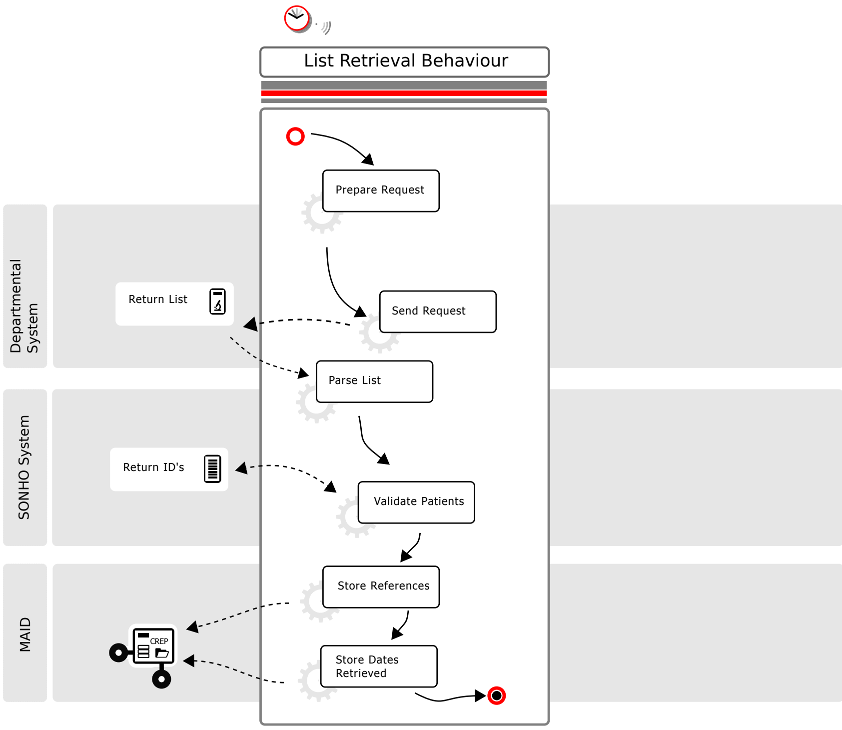


Figure 13: List Agent main actions and interactions during the report's references retrieval process.

of these reports a reference is obtained and stored in the CRep database. The behaviour is ciclic with a frequency based on timer events. Upon wake up the behaviour will determine the new interval based on the elapsed time since the last request and sends a new request to the DIS.

As lists of report references are retrieved, they are parsed, and the patient identification provided by the DIS is validated against the Administrative Discharge Transfer system database (SONHO). Like mentioned before, the system uses *num_sequencial* as the main patient identifier. In some cases this number is not present, the alternative numbers are used in order to find, in the central database, the *num_sequencial* which is used for clearly identifying the patient. When this number cannot be accurately identified an alert is produced and sent to the head of the department that produced the report.

Additionally, the time interval of the last successful reply is stored for each DIS and used when MAID is restarted (e.g. after a system reboot or crash). In this case, List agents adopt, concurrently, a recovery behaviour, that uses the last stored time interval, to recover the list of reports references produced during MAID's inactivity.

In Figure 13 are described the main actions and interactions that the List Agent undertakes to complete a report reference list retrieval process.

List agent interacts with the remote department by exchanging XML messages. The interaction is triggered by the List Agent by sending a 'request for reports' message that will produce a response including a list of reports references. Each request is identified by an element composed of two attributes: a unique identification and an time interval (Listing 1). The first is used for message management and control while the second attribute is used for selecting clinical reports generated during a specific period of time. Depending on the nature of the problem, unsuccessful requests are stored for a later request (e.g. network unavailability) or audit action (e.g. database unavailability).

Listing 1: Report List request example.

```xml
<!DOCTYPE MensagemHSJXXI SYSTEM "MensagemHSJXXI_4.dtd">
<MensagemHSJXXI id="1295644627.5" data-emissao="24/03/2009 11:43:36" versao-dtd="4">
  <pedido emissor="13" pedido="1295644627.5">
    <listaRelatorios data-inicial="18-03-2009 11:35:23" data-final="18-03-2009 11:42:25"/>
  </pedido>
</MensagemHSJXXI>
```

Each DIS reply is composed of a management element (including the request unique identification) and a list element containing references to reports produced during the requested time interval. Each report reference element includes, among other attributes, patient identification, author, type and retrieval location (URL) for the report.

Listing 2: Report references list example.

```xml
<MensagemHSJXXI id="0001234639388.5" data-emissao="01-09-2014 11:27:32" versao-dtd="4">
  <resposta emissor="7" pedido="1234639398.5">
    <listaRelatorios data-inicial="01-09-2014 10:22:03" data-final="01-09-2014 11:27:03">
      <relatorio id="1733211" tipo="Resultados de Exames (ImunoHemoterapia)" data-emissao="
          01/09/2014 08:20:00" operacao="Actualizar" genetica="0">
        <doente data-nascimento="26/03/1997" sexo="M">
          <nome>Anonymous</nome>
          <identificacao codigo="123456" dicionario="SONHO.NUM_SEQUENCIAL" />
          <identificacao codigo="654321" dicionario="SONHO.NUM_PROCESSO" />
        </doente>
        <episodio id="999999" tipo="I" />
        <listaIntervencoes>
          <intervencao funcao="Requisição">
            <pessoa>
              <identificacao codigo="DocID123456" dicionario="SIBAS.COD_MEDICO" />
            </pessoa>
            <servico>
              <nome />
              <identificacao codigo="DepID123456" dicionario="SIBAS.COD_SERVICO" />
            </servico>
          </intervencao>
        </listaIntervencoes>
        <listaDocumentos>
          <documento tipo="PDF">http://webgle:88/DadosSibas/GetDoc.aspx?doc=1733211</documento>
        </listaDocumentos>
        <HL7_Entry>
          <organizer classCode="CLUSTER" moodCode="EVN">
            <statusCode code="completed" />
            <effectiveTime value="20140901102627" />
            <component typeCode="COMP">
              <organizer classCode="BATTERY" moodCode="EVN">
                <code code="4" codeSystemName="SIBAS.GrupoExame" displayName="HEMOSTASE" />
                <statusCode code="completed" />
                <effectiveTime value="20140901102627" />
              </organizer>
            </component>
          </organizer>
        </HL7_Entry>
      </relatorio>
    </listaRelatorios>
  </resposta>
</MensagemHSJXXI>
```

The system evolved trough 4 versions of messaging format. The last version included an HL7 compatible section and a marker for the presence of genetic information in the report thus enabling enhanced specificity while enforcing access control by the main web interface.

An example of the list request and list request reply is depicted in Listing 1 and Listing 2.

4.2.3 *Report File Retrieval Process*

The second stage of making a report available is the retrieval of reports from the stored report references, this is accomplished by two types of agents the Balancer Agent, File Agent and Express Agent. Their model and main characteristics are described in the following subsections.

4.2.3.1 *Balancer Agents Model*

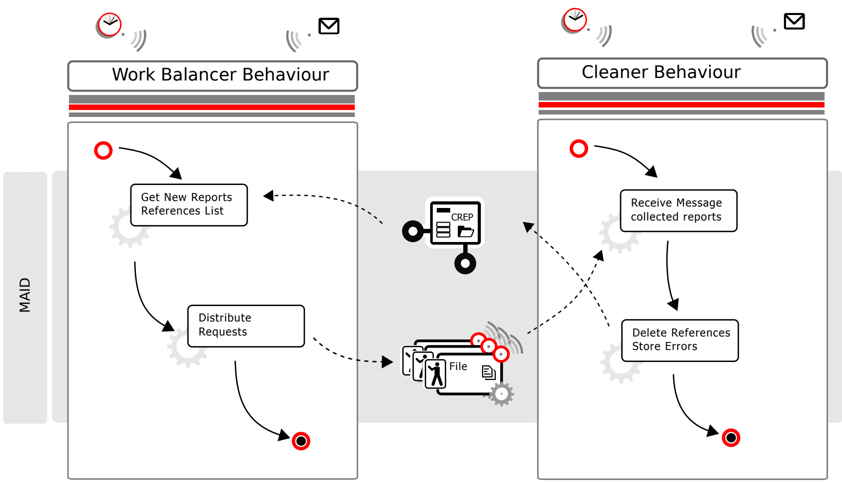


Figure 14: Balancer Agent main actions and interactions.

To each of the DIS is assigned an individual Balancer Agent. These agents regularly survey the CRep database looking for new clinical reports references stored by the List Agent. When new references are available, the agent collects them from the database and stores them in the internal pile. The references are described by an identifier and a URL for document retrieval. These references are distributed by a set of File Agents having in consideration their workload that can vary if some reports take more time to retrieve than others.

The Balancer Agent keeps track of what is distributed and to what agent it was distributed. It receives status messages from the File Agents regarding the requested references retrieval managing its internal list accordingly (eliminating the ones retrieved and logging the status of the ones that weren't retrieved). A schematic version of the main actions and interactions is depicted in Figure 14.

Messages sent to File Agents include a payload composed by a descriptive header with id for the request followed by a series of report references entries to be retrieved.

4.2.3.2 *File Agents Model*

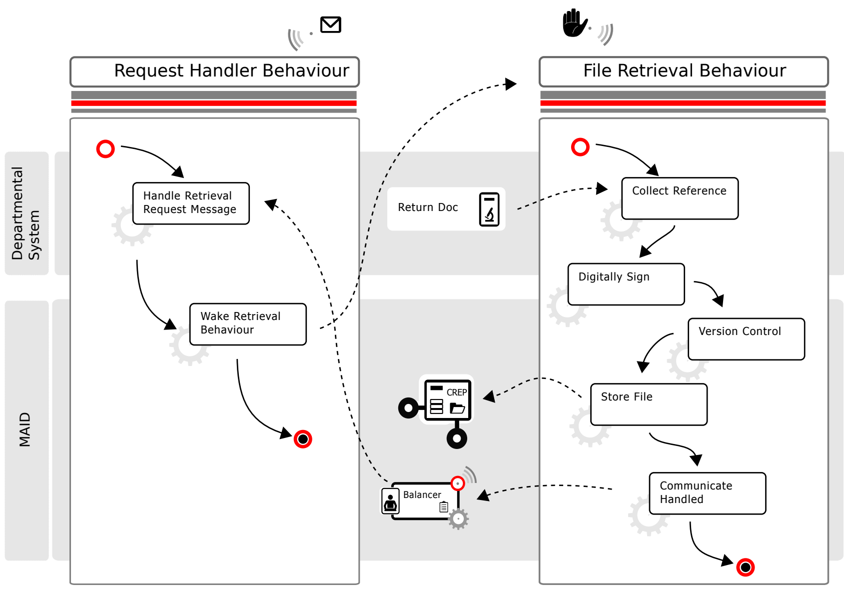


Figure 15: File Agent main actions and interactions.

As described above, the set of File Agents work under a Balancer Agent coordination. File agents are designed to accept a list of report references and to retrieve the associated documents from the corresponding DIS. The list is processed sequentially as reports files are requested to the DIS and their status is maintained according to success or unsuccess of the reference retrieval.

File Agent will store the document retrieved in the CRep filesystem. For each single patient there is a directory on the CRep File System. The directory path is determined by splitting the last two pairs of digits of internal patient identification number and concatenating the entire number in the end (e.g. internal patient number 123456789 is stored at /67/89/123456789/). This operation provides an uniform distribution on the directory tree. After determining the patient's directory, file agents retrieve the corresponding report using available network plug-ins. The report is then digitally signed and stored. Version control of the reports is performed using as fingerprint the SHA1 digest [179] of the location attribute (URL). A new version of a particular report makes the older version unavailable for users, though securely preserved for auditing purposes. A schematic version is depicted in Figure 15.

Upon collecting the reports a message is sent back to the Balancer reporting the status of the reports collected.

Messages sent by Balancer Agents to File Agents are composed by a descriptive header with id for the request followed by a series of report references entries (id and URL) to be retrieved. Messages received by the agent include the identification of the report and the status of the retrieval.

4.2.3.3 *Express Agent Model*

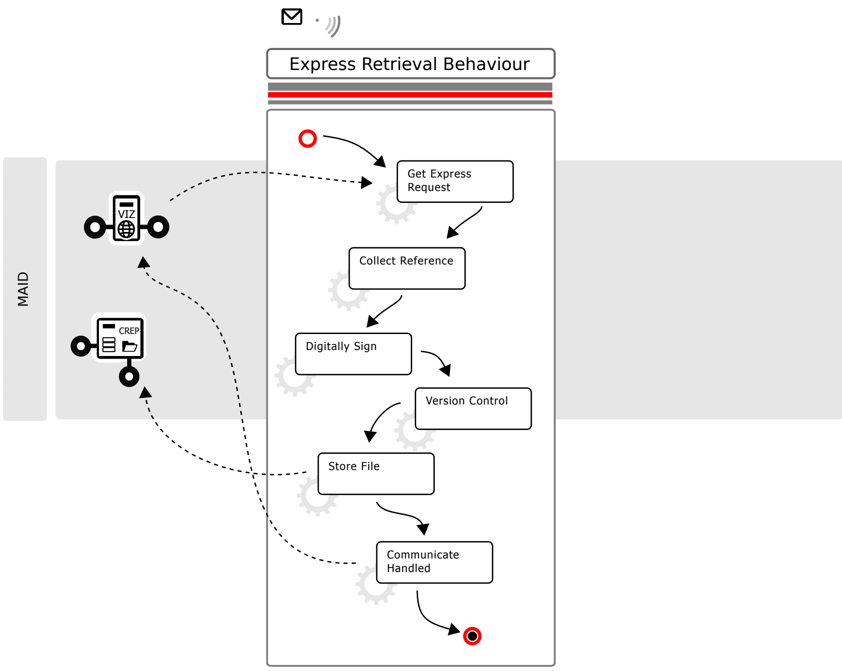


Figure 16: Express Agent main actions and interactions.

The express agent (Figure 16) was designed to act in response to request of document references that are still in the process of retrieval. As references are retrieved by List Agents they are made available right away to the user through the ICUViz interface. However, it can happen that these references are still unavailable on the CRep file system. This may happen when the report reference is in the process of being collected by the report file retrieval action.

Messages received by the Express Agent from the VIZ interface includes the patient identification and the report reference to be retrieved. According to the success of the retrieval, an acknowledgment regarding the readiness of the report is sent back.

4.3 USER INTERFACE

The departmental systems collected reports are made available through VIZ, an integrated web interface available throughout the hospital network. As soon as a reference list is retrieved, report references are made available to the end users aggregated in a list associated to the respective department. When the reports files are made available (either by the report collection process or by the express report retrieval process) they can be visualised by the health professionals within the system. The main interface, represented in Figure 17, is composed by the list of reports gathered for a given patient which also contains, at the bottom, a chronological bar constructed with the events date from SONHO. The selected individual reports are visualised within the interface, in Figure 18 are shown two examples of reports retrieved by the system.

Figure 17: User interface for report list view. It is composed by an patient search area at the top, the list of available reports in the middle and a chronological bar at the bottom with the past hospital encounters.

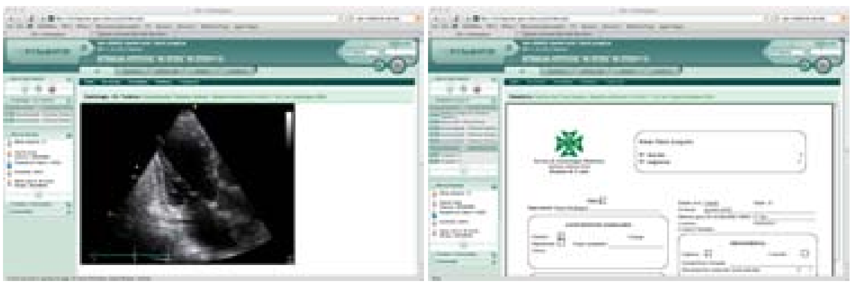


Figure 18: User interface for report document view, examples of an image exam and a discharge letter.

4.4 SECURITY

In order to address security requisites several mechanisms where developed and put in place.

For integrity enforcing and providing trust regarding the patient information stored all documents retrieved are digitally signed. When accessed, if the digital signature does not match with report contents then the report is marked as not valid.

Availability focuses on means to provide for the continuous access to information by authorised users. For this purpose equipment and power redundancy, backups and system monitoring were all put in place thus guarantying the system availability at all times and with the minimum downtime possible when errors occur. Additionally, any interruption of the system will trigger the recovery behaviour associated with the list collection process in order to guaranty that the produced reports are not lost.

The number of reports daily retrieved from each DIS is compared to what is expected based on previous departmental production and any deviation from expected values triggers an alert message to the system administrator.

The web interface includes authentication mechanisms, a role based access control mechanism and auditing tools that are used to give access, to control what access is given and to monitor users activity respectively.

4.5 DEPLOYMENT

MAID is implemented using JADE framework and its API. The system agents are deployed within a single container.

The system was firstly deployed in a Pentium 4 (1.8GHz), with 768Mb RAM and a Linux operating system. The central repository file system, which contains the clinical reports files, is located on a HP StorageWorks SAN, which is mounted in the MAID server using the NFS protocol. The database, which contains the patient's identification and references to the clinical records, is stored in two HP Server RP5740 RISC computer cluster running an Oracle v.9 database management system. In the last five years it has collected and stored more than 10000000 documents.

4.6 REPORTS REFERENCE LIST COLLECTION OPTIMISATION

From the analysis of the running system and looking at the rate of reports production, it was possible to detect patterns on DIS activity. It was observed a common variability in the production of reports throughout the day and through the days particularly on the weekends. Figure 19 represents patterns corresponding to two departments, and it shows the differences that can occur both within a department and between departments.

The observed patterns suggested the possibility of evolving MAID's agents static action scheduling to an adaptive behaviour that could reflect each system individual daily rate of report production. This could provide a decrease in the time elapsed between reports production and reports availability while trying to keep the departmental systems querying to the moments where reports are produced.

4.6.1 *Simulation Environment*

In order to test different approaches to the scheduling of agents actions and compare execution outcomes, it was necessary to recreate as much as possible

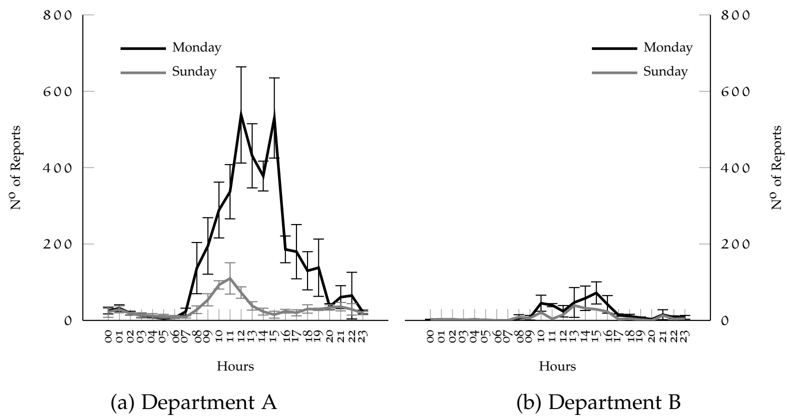


Figure 19: One month average report production of two departments.

the real scenario. For this purpose, a simulation environment was developed according to the general model depicted in Figure 20. It includes a replica of the repository (filesystem and database structure), a set of departmental information systems emulators and a simulation configuration interface.

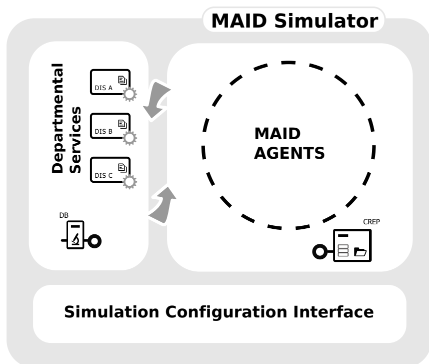


Figure 20: MAID Simulation Environment.

The departmental information systems replica emulates the real scenario report production by making use of a set of real data from the running system properly anonymized (only reports metadata and not the real documents themselves). For identifying the relevant reports for each request and returning the reference list it is used the date of emission from reports in the database repository. These are used to build the XML reply with the report references that each agent receives upon each report list request.

The simulation interface allows to edit each agent initial configuration files, to edit the simulation parameters (e.g. period of data to consider, log files path), to control the simulation execution and to edit the scheduling algorithm parameters.

4.6.1.1 *Changes to base system*

Some changes to the MAID agents were necessary in order to control the simulation flow. The Control Agent, already present in MAID for startup actions was extended to include an additional simulation related behaviour. This behaviour enables system's execution monitoring and other agents status gathering. At the end of the simulation, this agent stores the collected data into the database for further processing and analysis.

The List Agents were instrumented to be aware of the simulation starting date, using it as a baseline for calculating the necessary time adjustments to the date used for requests.

4.6.2 *List Agent Activity Scheduling*

In order to provide MAID list retrieval process with an adaptive action scheduling it was added to MAID a new Scheduler Agent. This agent is in charge of creating and providing weekly schedules for List Agent's actions by iterating through the list of active departments and reading the schedule model configuration files. The agent will then generate a weekly schedule map for each List Agent by processing the past history of report production and, according to a scheduling model determining, for each hour of each day of the week, the appropriate frequency cycle value. The weekly scheduling map is stored in the database in order to guaranty that in case of system execution interruption a normal resume of actions make take place without the need to recalculate the frequency cycle values again. List Agents will use these maps for managing their cycle of activities.

4.6.2.1 *Adaptive Cycle Frequency Scheduling Model*

The model proposed for determining an adaptive scheduling is based on the relationship between the higher cycle frequency value and maximum number of reports and the lower cycle frequency value and minimum number of reports determined from prior report production history. The line crossing these two points (Figure 21) would provide an approximation of the variation of cycle frequency value for a certain amount of expected report production.

$$\begin{cases} C_{max} &= a \times NR_{min} + b \\ C_{min} &= a \times NR_{max} + b \end{cases}$$

$$\begin{cases} a &= C_{max} - NR_{max} \times \dfrac{(C_{min} - C_{max})}{(NR_{min} - NR_{max})} \\ b &= \dfrac{(C_{min} - C_{max})}{(NR_{min} - NR_{max})} \end{cases} \quad (1)$$

The equation for this line is derived from the system of linear equations depicted in Equation 1 where:

C_{max} AND C_{min}: represent the predefined maximum and minimum allowed cycle frequency values, C_{max} can be also defined as the maximum time

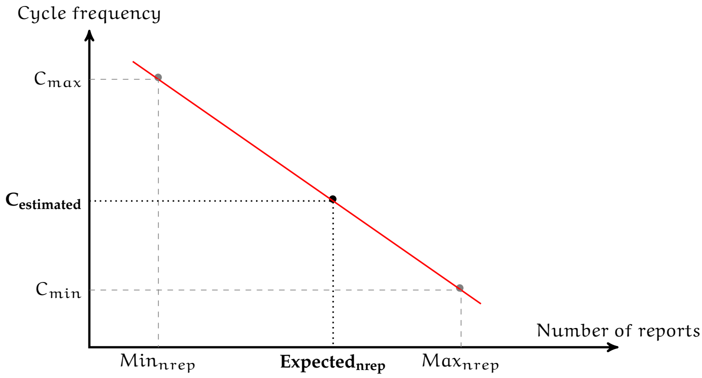


Figure 21: Report collection cycle frequency estimation.

elapsed from report production that is considered reasonable for it to be available (may vary between departments).

NR_{max} AND NR_{min}: the maximum and minimum number of reports produced for a given period (e.g. day, week).

The estimated frequency cycle values are derived from:

$$C_{estimated} = a \times NR_{Expected} + b \tag{2}$$

where the components a and b result from Equation 1 and

$C_{estimated}$: represents the collecting cycle frequency value for a given hour of the day of the week;

$NR_{expected}$: represents the expected number of reports for a given hour of the day of the week;

For the estimation of NR_{max}, NR_{min} and $NR_{expected}$, a central tendency measure (e.g. average, median) of past report production history can be used for reflecting the usual patterns of report production.

The period considered for NR_{max}, NR_{min} will determine if the daily schedule is influenced by the daily amount of report production (by determining the maximum and minimum of the all day) or by the overall amount of report production of the all week (by determining the maximum and minimum of the all week).

Regarding the $NR_{expected}$ it also can be influenced by the amount of past history that is used, either attaining a more stable nature from a median value for the same hour on the same days from the past month (eventually for the all year) or a more reactive nature if choosing the previous week report production values thus reflecting faster changes in the rates of production.

The cycle frequency value falls between the predefined C_{max} and C_{min} except when $NR_{expected}$ value is higher or lower than NR_{max} or NR_{min} respectively (which can happen if a big increase or decrease in report production happens). In these cases the values are forced to be bounded by C_{max} and C_{min}.

This approach allows the calculation of an adaptive cycle frequency value for each hour of the day of each week day, bounded by a pre-defined lower and

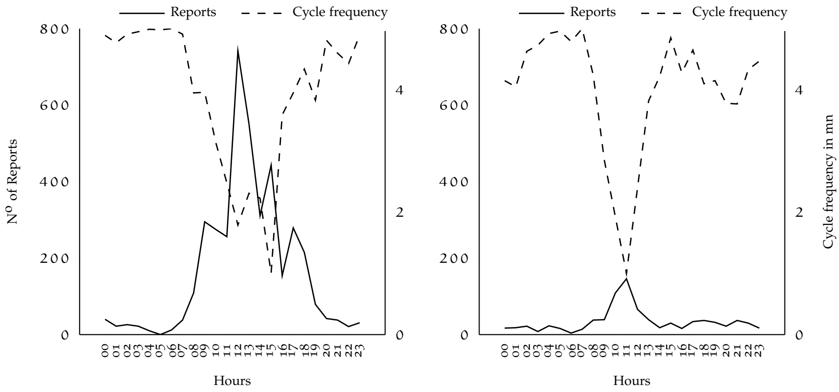


Figure 22: Cycle frequency values for Monday and Sunday with estimation of NR_{max} and NR_{min} from day data resulting in increased cycle frequency values even for a low amount of report production.

upper sealing, based on the past history rate of report production. An example of these scheduling is depicted in Figure 22 (with estimation of NR_{max} and NR_{min} from day data) and Figure 23 (with estimation of NR_{max} and NR_{min} from week data).

4.6.2.2 *Adaptive Cycle Frequency Scheduling Results*

The scheduling system was tested with two adaptive modes with estimation of parameters described in Equation 3 and Equation 4 and also with two static modes (5 mn and 1 mn cycle frequency values) for comparison. For testing purposes it was chosen a day of the week with the highest (monday) and lowest (sunday) correlation value with the mean of a month of reports.

$$\begin{aligned} NR_{max} &= Max(WMean(NR_{1pwd}, NR_{2pwd}, NR_{3pwd}) \\ NR_{min} &= Min(WMean(NR_{1pwd}, NR_{2pwd}, NR_{3pwd}) \\ NR_{Expected} &= WMean(NR_{1pwdh}, NR_{2pwdh}, NR_{3pwdh}) \end{aligned} \tag{3}$$

where NR_{npwd} represents the hourly number of reports of a particular day from the nth previous week. NR_{npwdh} represents the number of reports from the particular hour of a particular day of the nth previous week.

$$\begin{aligned} NR_{max} &= Max(WMean(NR_{1pw}, NR_{2pw}, NR_{3pw}) \\ NR_{min} &= Min(WMean(NR_{1pw}, NR_{2pw}, NR_{3pw}) \\ NR_{Expected} &= WMean(NR_{1pwdh}, NR_{2pwdh}, NR_{3pwdh}) \end{aligned} \tag{4}$$

where NR_{npw} represents the hourly number of reports for all days of the nth previous week. NR_{npwdh} represents the number of reports from the particular hour of a particular day of the nth previous week. $WMean$ represents the

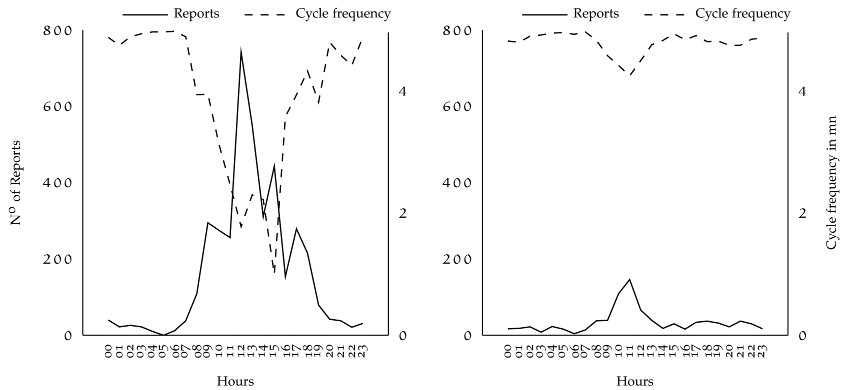


Figure 23: Cycle frequency values for Monday and Sunday with estimation of NR_{max} and NR_{min} from week data resulting in increased cycle frequency values only when the report production is high.

weighted mean using the weights 0.4;0.4;0.2 for the first, second and third previous weeks respectively.

Table 16: Average time elapsed from report emission to report collection for each model (in minutes) - Monday.

	24h Period (n=3887)	8-20h Period (n=3712)	07-21-23h Period (n=175)
	med (q1/q3)	*med (q1/q3)*	*med (q1/q3)*
$SM_{5\ mn}$	2:44 (2:00/3:40)	2:44 (2:01/3:40)	2:40 (1:43/3:47)
AM_{daily}	1:13 (0:45/1:57)	1:11 (0:44/1:52)	2:38 (1:26/3:34)
AM_{weekly}	1:23 (0:51/2:00)	1:21 (0:51/1:56)	2:28 (1:29/3:29)
$SM_{1\ mn}$	0:34 (0:24/0:45)	0:34 (0:24/0:45)	0:31 (0:20/0:41)

The median and inter quartil values regarding the elapsed time since date of report production (emission date) and the date of collection by the system (entry date) was evaluated and are presented in Table 16 for a Monday and in Table 17 for a Sunday. There are values for the all day, and in order to analyse the differences between more active periods against more inactive ones, values for the 8-20 period and for the 0-7 / 21-23 period. Table 18 and Table 19 represents the comparison of the median and interquartil values for differences between models regarding elapsed time from report emission to report collection.

It is possible to observe that the time for a report to be available when compared to the $SM_{5\ mn}$, decreases for the adaptive models and that this tendency is stronger in the 8/20 period whereas in the 0-7 / 21-23 period the adaptive models are closer to the $SM_{5\ mn}$ values. For Sunday the AM_{weekly} will be closer to the $SM_{5\ mn}$. This behaviour is further evidenced in Table 18 and Table 19 where the average difference between models regarding elapsed time from report emission to report collection are presented.

Table 17: Average time elapsed from report emission to report collection for each model (in minutes) - Sunday.

	24h Period (n=3887) *med (q1/q3)*	8-20h Period (n=3712) *med (q1/q3)*	07-21-23h Period (n=175) *med (q1/q3)*
$SM_{5\ mn}$	2:30 (1:21/3:48)	2:34 (1:25/3:50)	2:03 (0:57/3:22)
AM_{daily}	1:13 (0:36/2:14)	1:08 (0:35/2:06)	2:00 (0:58/3:20)
AM_{weekly}	2:25 (1:15/3:36)	2:24 (1:16/3:30)	2:39 (1:11/4:08)
$SM_{1\ mn}$	0:30 (0:15/0:47)	0:30 (0:15/0:47)	0:32 (0:15/0:47)

Table 18: Variation of elapsed time from report emission to report collection between models (in minutes) - Monday.

	24 h Period (n=3887)		8-20 h Period (n=3712)		0-7/21-23 h Period (n=175)	
	med (q1/q3)	*p*	*med (q1/q3)*	*p*	*med (q1/q3)*	*p*
AM_{daily} - $SM_{5\ mn}$	-1:28 (-2:22/-0:20)	*	-1:32 (-2:23/-0:24)	*	-0:21 (-1:19/1:05)	0.279
AM_{weekly} - $SM_{5\ mn}$	-1:19 (-2:18/-0:10)	*	-1:22 (-2:19/-0:13)	*	-0:25 (-1:04/1:01)	0.309
$SM_{1\ mn}$ - $SM_{5\ mn}$	-2:14 (-3:03/-1:25)	*	-2:14 (-3:02/-1:26)	*	-2:10 (-3:14/-1:11)	*
$SM_{1\ mn}$ - AM_{daily}	-0:40 (-1:25/-0:14)	*	-0:38 (-1:20/-0:13)	*	-2:08 (-3:08/-1:14)	*
$SM_{1\ mn}$ - AM_{weekly}	-0:51 (-1:29/-0:16)	*	-0:48 (-1:26/-0:16)	*	-1:59 (-3:00/-1:02)	*
AM_{daily} - AM_{weekly}	-0:08 (-0:19/0:30)	*	-0:09 (-0:19/0:30)	*	0:21 (-0:23/0:35)	0.217

* $p<0.0001$, Wilcoxon signed-rank test.

The adaptive models sit between the $SM_{5\ mn}$ and $SM_{1\ mn}$ models particularly in the 8/20 period while in the 0-7 / 21-23 period the adaptive models are closer to the $SM_{5\ mn}$ values mainly on Monday while on Sunday the AM_{weekly} will not differ very much from the $SM_{5\ mn}$. The differences in median time taken to collect reports between models have statistical significance and the tendency towards time reduction is evidenced in the 24h period and the 8/20 period for Monday and Sunday (except for the AM_{weekly} that is closer to the $SM_{5\ mn}$). The 0-7 / 21-23 period showed a lower divergence from adaptive models to $SM_{5\ mn}$ model both on Monday and Sunday and also between adaptive models on Monday.

Table 19: Variation of elapsed time from report emission to report collection between models (in minutes) - Sunday.

	24 h Period (n=830)		8-20 h Period (n=624)		0-7/21-23 h Period (n=206)	
	med (q1/q3)	*p*	*med (q1/q3)*	*p*	*med (q1/q3)*	*p*
AM_{daily} - $SM_{5\ mn}$	-0:54 (-2:28/0:25)	*	-1:03 (-2:41/0:25)	*	-0:08 (-0:27/0:48)	0.066
AM_{weekly} - $SM_{5\ mn}$	-0:06 (-1:37/1:27)	0.135	-0:07 (-1:40/1:21)	0.001	0:01 (-1:26/2:16)	0.23
$SM_{1\ mn}$ - $SM_{5\ mn}$	-2:01 (-3:18/-0:45)	*	-2:09 (-3:19/-0:55)	*	-1:38 (-2:45/-0:28)	*
$SM_{1\ mn}$ - AM_{daily}	-0:42 (-1:48/-0:12)	*	-0:37 (-1:37/-0:11)	*	-1:20 (-2:36/-0:32)	*
$SM_{1\ mn}$ - AM_{weekly}	-1:55 (-3:04/-0:47)	*	-1:55 (-3:00/-0:48)	*	-2:05 (-3:44/-0:27)	*
AM_{daily} - AM_{weekly}	-1:00 (-2:20/0:24)	*	-1:02 (-2:19/0:16)	*	-0:20 (-2:21/0:58)	*

* $p<0.0001$, Wilcoxon signed-rank test.

For each model it was also evaluated the number of total requests, the median and interquartil values regarding reports number per request and the number of empty report requests. These are represented in Table 20 for Monday and in Table 21 for Sunday.

Table 20: Report list variation among three scheduling models - Monday

SCHEDULLING MODEL	REQUESTS	REPORTS PER REQUEST	EMPTY REQUESTS
	n	*med (q1/q3)*	n (%)
$SM_{5\ mn}$	284	3 (0/17)	102 (36)
$Adapt_{weekly}$	419	5 (0/13)	119 (28)
$Adapt_{daily}$	413	6 (0/14)	120 (29)
$SM_{1\ mn}$	1415	0 (0/3)	775 (55)

Table 21: Report list variation among three scheduling models - Sunday

SCHEDULLING MODEL	REQUESTS	REPORTS PER REQUEST	EMPTY REQUESTS
	n	*med (q1/q3)*	n (%)
$SM_{5\ mn}$	283	1 (0/4)	128 (45)
$Adapt_{weekly}$	299	1 (0/4)	136 (45)
$Adapt_{daily}$	410	1 (0/4)	191 (47)
$SM_{1\ mn}$	1416	0 (0/1)	1052 (74)

The number of requests increases as the frequency cycle have lower values and consequently the number of empty requests gets higher. This is particularly evident in the $SM_{1\ mn}$ model where 55% on Monday and 75% on Sunday, of the total number of requests, did not include any report.

With the adaptive models it is possible to lower the elapsed time between report production and collection without incurring in highly inefficient collection actions.

The AM_{daily} model would be a more interesting approach given that it will reduce the time difference between report production and collection both in high and low report production rate days without increasing too much the number of requests.

4.7 DISCUSSION

As shown in the figure, MAID covers an area of the hospital information flow that was assured by at least one staff members of each laboratory plus another administrative on each of the main hospital clinical departments or hospital archives.

The use of multi-agent technologies allowed the successful implementation of an integration system for a large amount of heterogeneous clinical data. This patient information can now be accessed from any workstation in the hospital intranet thus making the implementation of a VEPR a useful reality. Also the system provides a user friendly and uniform view of clinical reports instead of the previous multiple interfaces from the various DIS, which were likely to confuse the user. Independence between DIS and central systems were guaranteed, enabling eventual upgrades of DIS to be made without need for major changes in the VEPR. No changes were required to local DIS hardware and full accessibility to previous collected patient data was maintained at all times.

As minor changes were required to the existing DIS, the implementation and deployment were done over a relatively short period of time and the stress in the organisation was very low. During all phases of the project the DIS didn't had to be interrupted. There was also no disruption in the clinical information as old records were kept. All these contributed to the relatively low financial costs. The process of integration of heterogeneous clinical information systems has shown the existence of several organisational or technical problems and, indirectly, contributed to its solution. While some reports cannot be associated with identified hospital patients (e.g. outpatients that are not administratively considered as hospital patients), some patients had multiple rather than unique identification numbers, making difficult their correct identification. A similar problem was found with staff identification numbers, which were reused after staff members left the hospital. Furthermore, hospital servers had major differences in their internal clock times, amounting to 25 minutes in some extreme cases. Therefore, use of the Network Time Protocol on every server involved in the VEPR project was mandatory. Also, at the time of development, some of the companies responsible for DIS lacked the skills to implement the necessary web-services for integration.

One of the major consequences of the implementation of MAID may be the reduction or even the elimination of administrative tasks related to the delivery and storage of clinical reports.

The development of a simulation environment provides for a safe testbed that will make possible to implement and test in a safely fashion new approaches to adaptive behaviour that may increase document availability time in real scenarios.

The process of collecting reports is prone to be implemented in an adaptive manner following the DIS report production rates. It can use the past history of document production from a Department to derive a activity scheduling that would reflect the need to accelerate or decelerate the rhythm of actions. With the adaptive model is possible to reduce the time necessary to make a report available adapting the rate of requests accordingly taking in consideration to only stress departmental systems when necessary. Given the time dependent nature of the frequency of actions it was difficult to accelerate the simulation process of reports collection which implied long running tests.

4.8 OUTCOMES

From the work described in this chapter resulted the publications described bellows. Besides these , a real world system was deployed and is in use in a major health institution in Portugal.

- Integration of hospital data using agent technologies - A case study. R. Cruz-Correia; P. Vieira-Marques ; P. Costa; A. Ferreira; E. Oliveira-Palhares; F. Araujo and A. Costa-Pereira. Ai Communications, Vol. 18, pages 191-200, 2005
- Maid - Multi Agent for the Integration of Data. P. Vieira-Marques, R. Cruz-Correia, P. Costa, E. Palhares, A. Ferreira, and A. Costa-Pereira. 1st Iberian Conference in Information Systems and Technologies (CISTI), Vol. I, pages 603–614, 2006.
- Simulation environment for the optimisation of the data retrieval capabilities of an agent based system in a healthcare setting. J.H. Patriarca-Almeida, P.M. Vieira-Marques, and R.J. Cruz-Correia. Holonic and Multi-Agent Systems for Manufacturing, Lecture Notes in Computer Science, Vol. 6867, pages 124-132, 2011 [5].
- Optimization of an agent based clinical data retrieval system. J.H. Patriarca-Almeida, P.M. Vieira-Marques, and R.J. Cruz-Correia. 7th Iberian Conference in Information Systems and Technologies (CISTI), pages 1–6, 2012.

5 PROFILING DATA NEEDS

In the previous chapter, a system for intra institution health data integration was described. However, most of the people tend to visit multiple health institutions during their lifetime raising the need for a broader data integration that goes beyond the institutional barrier. In this chapter, three studies are presented: the first one aims to geographically characterise the portuguese patient healthcare services usage patterns, the second aims to describe for a common clinical scenario what information is needed and the time restrictions that might be imposed in collection actions and the third aims to access if there is a variability in the importance attributed to clinical data in different contexts of healthcare by different types of health professionals.

5.1 PROFILING PATIENTS HEALTHCARE USAGE

5.1.1 *Objectives*

The objective of this work is to characterise patterns of geographic mobility of potential users of the National Health System. These movements may resulting from temporary or permanent displacements of population. This characterisation intends to assist in the identification of usage scenarios where an extended patient data access would be relevant and provide hints that could be used for guiding data search actions.

5.1.2 *Methods*

For population movements characterisation it was considered for analysis and interpretation two data sources:

1. INEP - Studies and statistical databases from the Portuguese National Institute of Statistics (INE) and from PORDATA portal (includes information provided by official entities, such as the Portuguese National Institute of Statistics and Eurostat)
2. PAR - Data from Patient Admission registries from all public hospitals in continental Portugal between 2000 and 2007

Data subjected to analysis was used to produce several tables and graphs in order to identify and highlight patterns of patient mobility.

The analysis of INEP data was aggregated geographically using NUTS (Nomenclature of Territorial Units) Type II [180].

The PAR data was aggregated by continental districts and for the first analysis, they were ordered from west to east and from north to south providing a representation that reflects approximately their relative distance. For additional analysis, data was aggregated based on the distances between the district capitals whose distance is above or below 100 km.

5.1.3 *Results*

5.1.3.1 *General population movements*

In Table 22, obtained from [181] are represented the numbers regarding population permanent movements for each region. These numbers reveal relevant mobility events in the resident population with a total amount of approximately 275000 people, (representing 3% of the total population), changing their permanent residency. These movements occur in all regions although with a higher presence in northern and centre regions (including Lisboa).

Inbound movements from foreign countries also have a reasonable expression accounting for 1% of the population and are more relevant in Norte and Lisboa regions.

Table 22: Population permanent movements in 2011 (based on residency in the previous year).

	TOTAL n	STAYED n (%)	FROM OTHER REGIONS n (%)	FROM OTHER COUNTRIES n (%)
Portugal	10562178	10087700 (96)	275615 (3)	81778 (1)
Norte	3689682	3550964 (96)	74279 (2)	25367 (1)
Centro	2327755	2,234783 (96)	53290 (2)	17012 (1)
Lisboa	2821876	2,656749 (94)	102827 (4)	26275 (1)
Alentejo	757302	725610 (96)	19504 (3)	4613 (1)
Algarve	451006	426868 (95)	13615 (3)	5168 (1)
Açores	246772	236387 (96)	5922 (2)	1146 (0)
Madeira	267785	256339 (96)	6178 (2)	2197 (1)

Besides the movements where there is an actual change of residence, the temporary displacement resulting from travel for leisure or business reasons represents also significative population movements. In these cases the actual use of health services may occur (e.g., accidents, sudden sickness) and most of the time, they will be provided by institutions that never had any contact with the patient.

Table 23: Travel according to reason and destination (in thousands).

	TOTAL	HOLIDAY	VISITS TO FAMILY	BUSINESS	HEALTH	RELIGION	OTHER
Total	17604.5	49.4	39.9	5.6	0.4	1	3.6
Norte	3595.4	38.9	48.6	7	0.3	1.2	4
Centro	5426.1	44.8	45	4	0.6	1.7	4
Lisboa	2871.7	35.7	50.2	6.7	0.9	0.4	6.1
Alentejo	2076.4	48.3	42.7	6.2	0.2	1.2	1.4
Algarve	3142.0	82.6	12.3	3.7	0	0.1	1.3
Açores	268.1	42.1	29.5	15.8	0.9	0.5	11.1
Madeira	224.8	64.1	20	14.3	0	0.2	1.5

In Table 23, it is represented the number of travel events, aggregated according to the reason for travel, having as destination each of the NUTS. These events are present in most of the regions with less expression in Madeira and Açores regions.

Besides the Business pretext, where most of the time the travel events may occur individually, the rest of the pretexts will eventually involve more than one person traveling together leading to a potential further increase in the numbers of population displacement.

5.1.3.2 *Displacement of population for health reasons*

The following data analysis uses data from discharge events of hospitalisations of patients according to their residence district and the health institution district where they occurred. In a first phase the data is analysed according to the place of treatment and in a second phase according to the origin of patients.

Table 24, shows the percentage of discharge events from patients who were hospitalised in each district, averaged from data running from 2000 to 2007, according to their district of origin.

Table 24: Percentage of resident and foreign patients treated in each district.

... comes from / Treated in...	V. Castelo	Braga	Vila Real	Bragança	Porto	Aveiro	Viseu	Guarda	Coimbra	C. Branco	Leiria	Santarém	Lisboa	Portalegre	Setúbal	Évora	Beja	Faro
V. Castelo	95.4	1	0	0	1.2	0.1	0	0	0.1	0	0	0	0	0	0	0	0	0
Braga	3.7	91.6	0.4	0.1	4.6	0.2	0	0	0.4	0	0	0	0.1	0	0	0	0	0.1
Vila Real	0	0.2	90.2	1.2	1.2	0.1	0.4	0	0.2	0	0	0	0	0	0	0	0	0
Bragança	0	0	2.6	97.7	0.8	0	0	0.1	0.2	0	0	0	0	0	0	0	0	0
Porto	0.6	7	2.9	0.5	86.1	2.2	0.2	0.1	0.7	0.1	0.1	0	0.1	0	0	0	0	0.2
Aveiro	0	0	0.1	0.1	4.3	96.2	0.2	0	12.9	0	0	0	0.1	0	0	0	0	0.1
Viseu	0	0	3.4	0.1	1.4	0.6	95	1.1	5.7	0.1	0	0	0.1	0	0	0	0	0
Guarda	0	0	0.1	0.1	0.1	0	3.5	96.1	4.2	2.4	0	0	0.1	0	0	0	0	0
Coimbra	0	0	0	0	0	0.3	0.2	0.3	52.9	0.4	0.2	0	0.1	0	0	0	0	0.1
C. Branco	0	0	0	0	0	0	0.1	1.6	3.8	95.8	0.2	1	0.2	0.3	0.1	0	0	0
Leiria	0	0	0	0	0	0.1	0	0.1	14.2	0.1	91.8	0.7	1.2	0	0.1	0.1	0	0.1
Santarém	0	0	0	0	0	0	0	0	3	0.2	5.7	92.6	2.8	0.2	0.3	0.2	0	0.1
Lisboa	0.2	0.1	0.2	0.2	0.1	0.1	0.4	0.3	0.8	0.5	1.9	1.1	87.4	0.4	2.5	0.4	0.3	0.8
Portalegre	0	0	0	0	0	0	0	0	0.2	0.2	0	4.3	0.6	90.3	0.2	3.4	0	0
Setúbal	0	0	0	0.1	0	0	0.1	0.1	0.2	0.1	0	0.1	4.6	0.2	95.2	0.7	2.9	0.3
Évora	0	0	0	0	0	0	0	0	0.1	0	0	0	0.8	8.4	0.6	91.9	0.3	0.1
Beja	0	0	0	0	0	0	0	0	0	0	0	0	0.7	0.1	0.6	2.8	95.9	0.4
Faro	0	0	0	0	0	0	0	0	0.3	0	0	0	1.1	0	0.3	0.1	0.4	97.8
Total (n)	20419	67967	24461	17263	174998	54866	32960	14466	89478	23873	34750	37446	235513	10449	64243	13380	11371	36060

From the analysis, as a function of the origin of patients admitted to each hospital, it is possible to observe that there is a greater predominance of admissions of patients from nearby districts and that there is an higher agglomeration around the hospitals of higher dimension (Porto, Coimbra and Lisboa). These are also the places with the highest percentage of total admissions: Lisboa (24%), Porto (18%), Coimbra (9%).

In almost all districts, the percentage of patients admitted are clearly from the same district ranging between 86% in Porto and 98% in Faro. The exception is Coimbra, where only 53% of the patients treated by local institutions are local patients. Of the remaining 47% admission events, it is possible to highlight the admissions of patients from Aveiro (13%) and Guarda (14%).

Table 25: Percentage of inpatient events where patient district of origin is >100km, <100km or local.

	Viana do Castelo	Braga	Vila Real	Bragança	Porto	Aveiro	Viseu	Guarda	Coimbra	Castelo Branco	Leiria	Santarém	Lisboa	Portalegre	Setúbal	Évora	Beja	Faro
<100	4.2	8.0	6.3	0.0	11.2	3.1	4.2	2.7	32.8	2.6	5.8	1.7	7.4	0.3	3.1	3.5	0.3	0.0
>100	0.4	0.4	3.5	2.3	2.6	0.7	0.8	1.2	14.3	1.6	2.4	5.7	5.2	9.5	1.7	4.5	3.8	2.2
local	95.4	91.6	90.2	97.7	86.1	96.2	95.0	96.1	52.9	95.8	91.8	92.6	87.4	90.3	95.2	91.9	95.9	97.8

In Table 25, it is described the percentage of patients that are admitted in a given district, grouped by their cities of origin, that are distant more than 100km. In this setting the districts that receive the most far away patients are from Center and South of Portugal, Coimbra (14.3%) have the highest percentage of patients that come from more than 100km away.

Table 26: Percentage of patients admitted in other districts' health institutions.

Patient from ... ... goes to	V. Castelo	Braga	Vila Real	Bragança	Porto	Aveiro	Viseu	Guarda	Coimbra	C. Branco	Leiria	Santarém	Lisboa	Portalegre	Setúbal	Évora	Beja	Faro	Total (n)
V. Castelo	86.8	3.1	0	0	9.1	0.1	0	0	0.5	0	0	0	0.2	0	0	0	0	0	22444
Braga	1	86.7	0.1	0	11.2	0.2	0	0	0.5	0	0	0	0.2	0	0	0	0	0	71830
Vila Real	0	0.5	88.8	0.8	8.2	0.1	0.5	0	0.7	0	0	0	0.2	0	0	0	0	0	24849
Bragança	0	0.1	3.3	87.5	7.7	0.1	0	0.1	0.9	0	0	0	0.3	0	0	0	0	0	19267
Porto	0.1	3	0.4	0.1	95	0.8	0	0	0.4	0	0	0	0.2	0	0	0	0	0	158697
Aveiro	0	0	0	0	10.5	73.1	0.1	0	15.9	0	0	0	0.3	0	0	0	0	0	72248
Viseu	0	0	2	0	6.2	0.8	77.2	0.4	12.6	0	0	0	0.6	0	0	0	0	0	40542
Guarda	0	0	0.1	0.1	0.6	0.1	5.8	70.3	19.2	2.9	0	0	0.8	0	0.1	0	0	0	19784
Coimbra	0	0	0	0	0.1	0.3	0.1	0.1	98.5	0.2	0.1	0	0.3	0	0.1	0	0	0	48070
C. Branco	0	0	0	0	0.1	0.1	0.1	0.8	12.4	82.7	0.2	1.4	1.9	0.1	0.1	0	0	0	27651
Leiria	0	0	0	0	0.1	0.1	0	0	26.5	0	66.3	0.5	6.1	0	0.1	0	0	0.1	48113
Santarém	0	0	0	0	0.1	0	0	0	5.9	0.1	4.2	74.9	14.1	0	0.4	0.1	0	0.1	46264
Lisboa	0	0	0	0	0.1	0	0.1	0	0.3	0.1	0.3	0.2	97.9	0	0.8	0	0	0.1	210345
Portalegre	0	0	0	0	0.1	0	0	0	1.4	0.4	0	12	10.6	71	0.9	3.5	0	0	13287
Setúbal	0	0	0	0	0.1	0	0	0	0.3	0	0	0.1	14.9	0	83.8	0.1	0.4	0.1	72916
Évora	0	0	0	0	0.1	0	0	0	0.5	0	0	0.1	11.7	5.6	2.6	79.1	0.2	0.1	15560
Beja	0	0	0	0	0.1	0	0	0	0.3	0	0	0	12.6	0.1	2.8	2.8	80.2	1.1	13215
Faro	0	0	0	0	0.1	0	0	0	0.6	0	0	0	6.6	0	0.5	0	0.1	91.9	38334

Table 26, shows the percentage of patients admitted in each district. From the analysis, as a function of the most common destinations of admissions in each district it is possible to observe a greater predominance, if the district of origin is excluded, of admissions in major centres Porto, Coimbra and Lisboa.

Every district receives patients from every district, although in many cases they are residual. It is possible to highlight some districts whose patient destinations are more intense like Leiria that have almost 35% of their patients treated elsewhere and the ones having more than 20% of the patients going to other districts health institutions (e.g., Aveiro; Viseu; Guarda; Leiria; Santarém; Portalegre; and Évora). Considering a barrier of 1%, the districts who have more different destinations are Portalegre and Beja with 4 different destinations.

In Table 27, it is possible to identify the districts who have more than 10% of their patients traveling more than 100km to be admitted. These are mainly inland districts: Portalegre(29%); Guarda (21%); Évora (18%); Beja (17%); Castelo Branco (16%) and Bragança (13%).

Table 27: Percentage of inpatient events where patient destination is >100km, <100km or local.

	Viana do Castelo	Braga	Vila Real	Bragança	Porto	Aveiro	Viseu	Guarda	Coimbra	Castelo Branco	Leiria	Santarém	Lisboa	Portalegre	Setúbal	Évora	Beja	Faro
<100	12.3	12.2	8.7	0.0	4.3	26.4	15.8	8.7	0.5	0.9	27.0	18.3	1.0	0.4	15.0	2.8	2.8	0.0
>100	0.9	1.1	2.5	12.5	0.7	0.5	7.0	21.0	1.0	16.4	6.7	6.8	1.1	28.6	1.2	18.1	17.0	8.1
Local	86.8	86.7	88.8	87.5	95.0	73.1	77.2	70.3	98.5	82.7	66.3	74.9	97.9	71.0	83.8	79.1	80.2	91.9

5.1.4 *Potential nancial savings*

From the total number of admissions and taking in consideration only 10% of the cases of "foreign" patients that are treated in a particular district and a set of exams that could be needed and reused, potentially, almost one million euro would be saved annually. The table Table 28 reflects distribution of these savings per district and the total amount.

Table 28: Potencial financial savings when data can be reused. Individual values taken from the portuguese National Health System price tables determined by the Portuguese Health Ministry and used for financing the health care institutions.

	UNIT VALUE	PORTO	COIMBRA	LISBOA	OTHERS	TOTAL
X-Ray	5	12123	21070	14811	13419	61423
ECG	6.5	15760	27391	19254	17445	79850
CT	8.7	21094	36662	25771	23349	106876
EchoCG with Döppler study	53.2	128991	224185	157586	142779	653541
Total		177968	309308	217422	196992	901690

5.1.5 *Discussion*

From the analysis of this data it is possible to observe that the population reveals relevant internal migratory movements with permanent changes in residency, along with temporary seasonal movements for business or leisure purposes. Additionally, and regarding health related reasons reflected by the admissions in hospitals around the country, it is possible to observe that the major cities reveal a strong affluence from every district but predominantly from nearby districts. The characterisation for most common destination for admissions of patients from each district keeps this centrality in the large centres. There is a high number of patients whose districts for admission are distant over 100 km.

These results also reflects the national organisation of hospitals of reference associated to each district where less resources are available (particularly inland institutions), it would be interesting to do this analysis based on emergency events and not only by the admissions that are somehow influenced in some cases by the "hospital of reference rule", unfortunately, it was not possible to obtain this additional data.

Nevertheless, this characterisation justifies the need for having information exchange, either because medical events can occur during travel, and usually there is no local information about the patient, but also because patients are in some cases treated outside their district reference hospitals.

The evidence gathered may provide hints for the development of information search algorithms that would be better adjusted to more probable locations for finding patient information. The amount and the dispersion of admission events suggests that in most of the districts, a district and regional health institutions search should be given priority followed by a broader scope of search.

5.2 IDENTIFYING RELEVANT DATA IN SELECTED MEDICAL CARE SITUATIONS

The purpose of this work was to identify and illustrate what kind of information is important, where to search and the time constraints that apply while gathering information according to a particular situation of medical care.

Given the most frequent attending reason in a central Hospital , a storyboard and use case were created for describing the main characteristics of the medical condition. The list of relevant clinical and diagnostic data was initially based on established guidelines ([182],[183]), and then adjusted according to interviews with specialists aiming to refine and validate the process flow and determine the relevancy and collection time frames of the information along the process of medical attending.

To illustrate the storyboard, a process chart was developed including the various stages of care provision that the patient undertakes. Integrated with health care events there is a representation of collection tasks and identification of types of data to be collected according to their priority. Another important element developed was the table describing what is considered important information and where it can be found. The patient clinical data can be largely spread along the different types of institutions (e.g. hospitals, primary care clinics, labs).

5.2.1 *Methods*

In order to choose a relevant medical condition to be studied, it was identified the most frequent admission reason of a large University Hospital (Hospital São João – Porto, Portugal). The list generated by the DRGs statistical analysis software name ARCHI, pointed out pregnancy and labor as the top reason, corresponding to 10,7% of the total admission reasons in 2009.

Due to the relevancy of the numbers (the second main reason was pneumonia, with 1,82%), it was decided to focus on the pregnancy and labor. Based on the defined theme, it was made a storyboard to serve as a keystone.

At last, a set of interviews was conducted with three gynaecology and obstetrics specialists (two Portuguese and one Brazilian) to define and validate the patient process flow, considering the various stages and their optimistic mean durations (according to the specialist experience). It was also asked the specialists to analyse the list of relevant clinical and diagnostic information, allowing him/her to suggest inclusions or exclusions, and to determine the priority relevancy of the information according to its potential use period along the usecase patient attending flow. So it would be possible to determine the priority order of the patient's information to be electronically collected at other health care facilities (e.g. hospitals, labs, health centres), the necessary time to accomplish the task and in manner to do it without consuming the ICT resources.

5.2.2 *Results*

The conducted interviews allowed the development and consolidation of the process flow, detailing the various events along with an optimistic estimated mean time duration (described in minutes inside each bar) and represented by the top bar of Figure 24.

Pregnancy emergency situation: *"A 29 years old female patient, from Vila Nova de Gaia, 38 weeks pregnant, was admitted in the Hospital São João Emergency Department (ED) due to abdominal pain and forgot her pregnancy book at home. Prenatal care was done in a primary care facility, a private practice and lab analysis in two different institutions."*

Figure 24: Usecase of an Emergency admission event with the Patient Information Flow: On the top is the patient flow bar, which describes the various events along a typical attending. Inside each bar is the duration (minutes) of the referred event. The other bars describe the process of collecting and making data available before a possible use (minutes are inside each bar).

The estimated mean time of data usage was an optimistic one because it should represent better a constrained limit to search and collect the health care data. An important point to clarify in the understanding of the patient bar is the workup stage. It is known by the health care professionals as the period when many necessary actions related to diagnosis and treatment are performed.

During the workup is when necessary medications are administrated and when, for example, blood is collected for analysis. At this time, there is also the participation of other important health professionals (e.g. nurses, technicians).

In Figure 24 the patient flow bar illustrates a possible example of an automated data collection process divided into two groups according to the priorities.

Table 29: Classification of relevant data and most probable sources.

		Patient	Hospitals	Health centres	Private clinics	Labs	Image clinics
Priority I Data	Allergies	⊙	⊙	⊙	⊙	⊖	⊖
	Blood analysis	⊙	⊕	⊙	⊙	⊕	⊖
	Blood pressure	⊙	⊕	⊕	⊕	⊖	⊖
	Blood type	⊙	⊕	⊕	⊕	⊕	⊖
	Corrected estimated due date	⊙	⊕	⊙	⊕	⊖	⊙
	Estimated due date	⊕	⊕	⊙	⊕	⊖	⊙
	First trimester USG report	⊙	⊕	⊙	⊙	⊖	⊕
	Group B streptococci test at ± 36 weeks	⊙	⊙	⊙	⊕	⊕	⊖
	Last menstrual period	⊕	⊙	⊙	⊕	⊖	⊙
	Obstetric history	⊕	⊕	⊕	⊕	⊖	⊖
	Second and third trimester USGs reports	⊙	⊙	⊙	⊕	⊖	⊕
Priority II Data	Chosen anesthesia	⊖	⊙	⊖	⊖	⊖	⊖
	Clinical records	⊖	⊕	⊙	⊙	⊖	⊖
	Contraception	⊕	⊙	⊙	⊙	⊖	⊖
	Depression (-)	⊙	⊙	⊖	⊙	⊖	⊖
	Domestic violence (-)	⊙	⊖	⊖	⊖	⊖	⊖
	Environment and lifestyle	⊕	⊙	⊙	⊙	⊖	⊖
	Fetal doppler	⊖	⊕	⊖	⊙	⊖	⊕
	Menstrual history	⊕	⊙	⊙	⊙	⊖	⊖
	Prenatal lab studies	⊙	⊙	⊙	⊙	⊕	⊖
	Previous pathologies	⊕	⊕	⊙	⊕	⊖	⊖
	Weight records	⊙	⊙	⊙	⊕	⊖	⊖

⊕ - Highly Probable ; ⊙ - Probable ; ⊖ - Not Probable

The Priority I data is the most required information for the early stage of health care attending. The Priority II data is also considered important, but it includes information that would be used by the health care professional at a later stage.

The colours indicates the availability of the data, red bars indicate unavailable data and green bars mean available data. The set of interviews also made possible to identify the relevant patient health information considering the selected use-case situation and to create a list of possible sources better suited to hold the information.

In Table 29 are described the Priority Data data groups and the main source of information. In order to better illustrate the relevance of each source, three probability levels were defined and are identified by the symbols $\oplus$ - Highly Probable ; $\odot$ - Probable ; $\ominus$ - Not Probable.

5.2.3 *Discussion*

The presented results can be useful when considering an automated search for health care information of a patient. As this work considered a situation related to an HED and emergency situations, there is considerably little time to make the information available to the health professional, hence the efforts should be optimised and directed to the more relevant information.

The first step would be to determine an estimated mean time until the first medical care moment (admission time). The previous events are most of them composed by waiting time, so it is crucial to adequate an automated search according to the mean time of the hospital to collect the required information. The importance of the optimistic estimated mean time can be better understood at this stage, once it is acceptable to have the information grouped before the medical care, but it is not acceptable if we consider the opposite order of the events.

Also, it was important to determine the Priority I Data to be collected within this interval. As the Priority II Data has a more extended interval, which includes the first medical care and the workup (counting with more waiting time), it allows to collect more and heavier information (e.g. image files). An interesting finding was to evidence the sources with most probability to hold information. Although the patients are considered to be the primary source of information, the Obstetrics and Gynaecology specialists ranked them as the third source. This is explained by the fact that the chosen intitutions are the ones holding the information, however when no information exchange system is in place, it is the patient who takes guard of the information and makes them available at the point of care.

The specialists considered the hospitals and the private clinics as the sources with greater probability to hold information, counting 10 $\oplus$ items. The patients were the following source with great probability with 7 $\oplus$ items.

Combining the estimated time of the events along the usecase patient attending flow, the information considered relevant by the specialists with a list of the most probable sources to hold the information, it is possible to design an adaptive collecting method thus optimising the efforts to obtain relevant data in a highly distributed scenario.

This characterisation can also help to identify the most relevant data to different professionals involved at each step of health care specially when time constraints exist. By identifying critical periods for data availability and providing input regarding the most important types of data at a given moment in the care process, basic functions such as agent activation, search and transport of clinical information may be highly improved. In this system, agent's goals maybe be directed at addressing group or individual agendas (heterogeneous health professionals and patients) regarding health information needs. Nevertheless, to achieve an efficient use of the resources the information to be searched, their relevance, their sources and the available time to accomplish the tasks must be personalised according to each medical situation.

5.3 PROFILING DATA NEEDS IN MULTIPLE CONTEXTS OF CARE

Health Information Systems (HIS) integration increases access to patient data and may improve quality of care. However, providing large amounts of unprocessed data may have the opposite effect. Few studies have explored whether there is a variance in importance attributed to different types of data in different healthcare settings. This work aims at characterising the access to external patient information usage profile and explore the variability in the importance given to access to different types of patient data(e.g., Alerts, diagnosis, treatments) and sources (e.g., hospitals, primary care centres, laboratories) access, in different types of healthcare encounters, by different medical doctors (MD) profiles.

It was also included an assessment of health professionals opinion regarding the importance and impact of having HIS integration.

5.3.1 *Objectives*

For the purpose of this characterisation the following study questions were defined:

1. What is, within the institution, the current status of access to main groups of clinical information?
2. What is the impact expected from access to external clinical data?
3. What are the main sources of information and their relevance in three scenarios of clinical practice (Emergency, Inpatient and Outpatient) and the place of work (Hospital and Primary Care)?
4. What are the main types of information and their relevance in three scenarios of clinical practice (Emergency, Inpatient and Outpatient) and the place of work (Hospital and Primary Care)?

5.3.2 *Methods*

STUDY DESIGN AND SETTING

The study sampling consisted on MDs from selected healthcare institutions (chosen by geographical proximity convenience) and attending post-graduation courses at Faculty of Medicine University of Porto.

In total 126 paper questionnaires were distributed to Medical Doctors (MD) during 2010; questions regarding expected impact of having access to external clinical data were adapted from Shapiro et al. ([184]).

This study consisted of two phases. On the first phase a set of questionnaires (n=15, response rate 100%) including only the section "Importance of data types in multiple contexts" was distributed. On the second phase, a set of extended questionnaires (n=111, response rate 64%) was distributed including the additional sections that are described in detail in the next sections.

QUESTIONNAIRES

The questionnaire is composed by closed, multiple-choice and scale questions and is included in Appendix A.

It starts with a **Characterisation** section regarding MDs demographic data and past clinical experience. Next, there is a section **Current Status** concerning the status of access to major groups of clinical information in the workplace. Following this section, there is an **Obstacles and Benefits** section focused on the perspective of the inquired regarding the main obstacles and benefits in accessing information proceeding from external institutions. Finally, the questionnaire ends with two sections: **Importance of Data Sources** and the **Importance of Data Types in multiple contexts**. In these, the MDs were asked to classify the relative importance of access to different types of data sources and to different types of patient data in three settings of care provision: Emergency, Inpatient and Outpatient.

DATA COLLECTION AND ANALYSIS

The questionnaires were hand held and followed, in the healthcare institutions, the formal validation protocol regarding ethics and departmental organisation hierarchy (Ethics Reference 89/2011).

Profiles for each encounter type (Outpatient, Inpatient and Emergency) were defined by grouping MDs who spend 70% or more of their time in one type of encounter. The remaining MDs whose time was distributed among the different types of encounters were grouped in an Undifferentiated Profile. This profile grouping was used to analyse the relative importance attributed to access to data sources and types of data (Objectives three and four) by different profiles of MDs.

For the classification of different types of data importance for each scenario, case answers were selected from the dominant profile scenario (discarding the answers given in other scenarios), for the Undifferentiated profile a median of all scenarios responses was used.

For variability analysis regarding the importance attributed to the different types of data by each profile, items from the same group of patient data were merged into a single answer by selecting the median from each of the dominant scenario answers.

For the analysis of variability in the importance given to exams and exam reports, cases used were those where reports was given higher importance than the exam itself in the dominant profile.

BIAS

A convenience sample can lead to a bias, however, as the objective of the study is to compare groups of MD profiles and the sample methodology was the same for all groups, we believe that the study aims could be achieved with this methodology.

STATISTICAL METHODS

Medians (Med) along with 25th and 75th percentile represented by (p25, p75) were used for describing variables. Chi-square test was used for comparing the percentage of patients where information access is tried and the main communication vehicle used for external access to information. Mann Whitney test was used for comparing the importance given to different types of patient data and data sources between MDs from Hospitals and Primary Care. Kruskal Wallis test was used for comparing the importance given to different types of patient data and data sources among the different profiles. For a comparison between percentages of MDs that give more importance to Complementary Exam Reports than to the Exams themselves in Outpatient and Emergency Profiles, Chi-square or exact-Fisher were used when applicable. A significance level of 5% was used throughout the analysis. Statistical analysis was conducted using SPSS 21.

5.3.3 *Results*

PARTICIPANTS: Respondents were 48% female and 52% male. Median of age of respondents is 39 years. 64 MDs worked in Hospitals (HMD) and 22 from Primary Care Centers (PCHC) represented as PCMD. 38% were classified with Outpatient profile, 38% as Emergency, 4% with Inpatient and 20% with Undifferentiated profile.

All respondents reported having access to a computer in their workplace.

Table 30: Absolute and relative frequency of Medical Doctors from hospital and from Primary care who claim that try to access to external sources and main communication interface used to obtain external information.

	INSTITUTIONS OF ORIGIN		
	Hospital n (%)	Primary Care n (%)	*p*
Access to external sources is tried in:			0.306
0-25 % of patients	27 (43)	2 (25)	
26-50 % of patients	20 (32)	5 (63)	
51-75 % of patients	9 (14)	0 (0)	
75-100 % of patients	7 (11)	1(13)	
External information is obtained by:			**0.006**
HIS	4 (7)	4 (50)	
e-Mail	1 (2)	0 (0)	
Phone	38 (64)	3 (38)	
Letter	16 (27)	1 (13)	

WHAT IS THE CURRENT STATUS OF ACCESS TO MAIN GROUPS OF CLINICAL INFORMATION?
From the results collected, 25% of HMDs try to obtain external information in more than half of their patients and most of the PCMDs only try to obtain external information in less than 50% of their patient encounters. However, we did

not found significant differences between the HMDs and PCMDs regarding the percentage of patients in which the access to external sources is tried (Table 30).

Significant differences regarding the main communication interface used to obtain external information were found between HMDs and PCMDs. Actually, most of the HMDs prefer phone, and most of PCMDs choose HIS (Table 30).

Table 31: Accessibility to main groups of information.

	HOSPITAL	PRIMARY CARE
	n (%)	n (%)
Administrative Data	60 (100)	8 (100)
Imagiology Results	62 (98)	0 (0)
Laboratory Results	61 (98)	1 (13)
Diagnosis	59 (95)	8 (100)
Emergency Episodes	59 (94)	3 (38)
Discharge Notes	57 (92)	4 (50)
Surgical Treatments	55 (90)	6 (75)
Pharmacological Treatments	53 (88)	8 (100)
Outpatient Episodes	53 (86)	8 (100)
Alerts	50 (86)	5 (63)
Surgical Reports	49 (82)	2 (25)
Inpatient Episodes	43 (74)	6 (75)
Day Care Episodes	42 (72)	2 (25)
Nursing Notes	41 (71)	1 (13)
Patient Managed Data	2 (4)	4 (50)

Results show that in Hospitals most of the main groups of clinical Information can be accessed through the local HIS with the exception of Patient Managed Data. In Primary Care, Lab Results and Nursing Notes are stated to not be so widely available through the local HIS (Table 31).

Main obstacles that were pointed out for external information accessibility are interoperability/integration issues in HIS (26%) followed by concerns regarding security/confidentiality (9%).

WHAT IS THE IMPACT EXPECTED FROM ACCESS TO EXTERNAL CLINICAL DATA?

Both HMDs and PCMDs classify as Very Important the access to patient data from other institutions; MDs classify as Difficult and PCMD as Somewhat Difficult the access to external data from their workplace (Table 32).

Regarding the classification of the impact of access to external information both MDs and PCMDs mentioned to Agree that this would benefit the ways of providing care, the department procedures and the institutional procedures.

Table 32: Impact from access to external information.

	TYPE OF INSTITUTION		
	HOSPITAL	PRIMARY CARE	
	Med(p25; p75)	Med (p25; p75)	*p*
Importance given to external data access [1]	5 (4,5)	5 (5,5)	0.246
Difficulty to access external data at your workplace [2]	5 (4,5)	4 (3,5)	0.146
Your access to external data would benefit [3]			
Way of providing care	2 (1,2)	2 (1,2)	0.725
Department	2 (1,2)	2 (1,3)	0.688
Institution	2 (1,2)	2 (1,2)	0.983
National Health System	1 (1,2)	2 (1,2)	0.344
Access to external data would increase/decrease [4]			
Efficiency	2 (1,2)	2 (1,2)	0.495
Number of test requested	4 (4,4)	4 (4,5)	0.502
Number of medical errors	4 (4,4)	4 (4,4)	0.463
Time available to the patient	4 (2,4)	4 (3,4)	0.685

[1] Scale: 1 (Unimportant) to 5 (Very Important).

[2] Scale: 1 (Easy) to 5 (Difficult).

[3] Scale: 1 (Totally Agree) to 5 (Totally Disagree).

[4] Scale: 1 (Largely Increase) to 5 (Largely Decrease).

MDs mentioned to Totally Agree that it would benefit the National Health Service in general while PCMD mentioned to Agree (Table 32).

Additionally, it is referred that this would increase the efficiency of their work; decrease the number of exams and test requests, the number of medical errors as well as the time spent for decision-making. It is relevant to note that there was no significant differences found between profiles.

WHAT ARE THE MAIN TYPES OF SOURCES OF INFORMATION AND ASSOCIATED RELEVANCE IN THREE SCENARIOS OF CLINICAL PRACTICE (EMERGENCY, INPATIENT AND OUTPATIENT) AND PLACE OF WORK (HOSPITAL AND PRIMARY CARE)?

When using MDs profiles for analysis, there are significant differences in the relevance given to data from Private Practice ($p=0.048$) among the MD's profiles. These differences are manifested by Emergency MDs who give more importance than other profiles to this type of data(Table 33).

Also, it is observable that the Outpatient profile gives more importance to Primary Care while the other profiles give more importance to Hospital sources (Table 33).

When MDs were asked to rank which sources of data (Hospitals, PC, Private Practice or Private Labs) are more relevant, HMDs ranked (from more relevant to less relevant): data from other Hospitals followed by data from PC, Private Practice and data from Private Labs. Whilst PCHC MDs ranked as more relevant data from other PCHC followed by data from Hospitals, Private Labs, and Private Practice.

Table 33: Importance given by medical doctors to different types of data sources according to profile type.

	TYPE OF PROFILE				
	OUTPATIENT	INPATIENT	EMERGENCY	UNDIFFERENTIATED	
SOURCES OF DATA[1]	Med(p25; p75)	Med(p25; p75)	Med(p25; p75)	Med(p25; p75)	*p*
Hospitals	3 (3; 4)	4 (3; 4)	4 (4; 4)	4 (3; 4)	0.658
Primary Care	4 (3; 4)	3 (3; 4)	3 (2; 3)	3 (3; 3)	0.114
Commercial Labs	2 (2; 2)	1 (1; 2)	1 (1; 3)	2 (1; 3)	0.427
Private Practice	1 (1; 2)	2 (1; 2)	2 (2; 3)	2 (1; 2)	**0.048**

[1] Scale: 1 (Low importance) to 4 (High Importance).

Significant differences were found regarding the importance attributed to hospital sources from HMD and PCMD (p=0.044) and regarding PC sources from HMD and PCMD (p=0.032) Table 34). Although, not significant, it is also observable an inversion on the importance attributed to Commercial Labs and Private Practices between HMDs and PCMDs.

Table 34: Importance given by medical doctors to different types of data sources according to institution type.

	TYPE OF INSTITUTION		
	HOSPITAL	PRIMARY CARE	
SOURCES OF DATA[1]	Med. (p25; p75)	Med. (p25; p75)	*p*
Hospitals	4 (3; 4)	3 (3; 3)	**0.044**
Primary Care	3 (3; 3)	4 (3; 4)	**0.032**
Commercial Labs	2 (2; 2)	1 (1; 2)	0.312
Private Practice	1 (1; 3)	2 (2; 2)	0.135

[1] Scale: 1 (Low importance) to 4 (High Importance).

WHAT ARE THE MAIN TYPES OF INFORMATION AND ASSOCIATED RELEVANCE IN THREE SCENARIOS OF CLINICAL PRACTICE?

Significant differences among different MDs profiles (p=0.003) were found regarding the importance given to chronic conditions (diabetes and allergies), in this case Inpatient MDs give less importance to this type of data. Also, a significant difference was found (p=0.038) regarding image exams where Inpatient MDs give less importance to this type of data than other profiles (Table 35).

When grouped by institution of origin, image exams are considered less important by PCMD than by HMD, with a significant difference (p=0.010). To note also that Patient Managed Data is more important to PCMD than to HMD (p=0.001) (Table 36).

When comparing the relative importance of an image exam and its report, Outpatient MDs give, in a higher percentage, higher importance to the exam

Table 35: Importance given by medical doctors to different types of data according to profile type.

	TYPE OF PROFILE				
	Outpatient	Inpatient	Emergency	Undifferentiated	
TYPES OF DATA[1]	Med(p25; p75)	Med(p25; p75)	Med(p25; p75)	Med(p25; p75)	*p*
Administrative Data	2 (1,2)	1 (0,2)	2 (1,2)	1.5 (1,2)	*0.265*
Alerts (allergies. etc.)	2 (2,2)	1.5 (1,2)	2 (2,2)	2 (2,2)	***0.003***
Diagnosis	2 (1.5,2)	1 (1,2)	2 (2,2)	2 (2,2)	*0.117*
Treatments	1.5 (1,2)	1.5 (1,1.5)	2 (1.5,2)	1.5 (1.4,2)	*0.275*
Clinical Events	2 (1,2)	2 (2,2)	2 (2,2)	2 (1.6,2)	*0.365*
Lab Exams	2 (2,2)	2 (2,2)	2 (1,2)	2 (1.5,2)	*0.378*
Image Exams	1.75 (1,2)	1 (1,2)	2 (1.5,2)	2 (1,2)	***0.038***
Image Exams Reports	2 (2,2)	2 (1,2)	2 (1.5,2)	2 (2,2)	*0.532*
Other Exams	2 (1,2)	2 (2,2)	2 (2,2)	2 (1.5,2)	*0.362*
Patient Mng. Doc.	2 (1,2)	1.5 (1,2)	1 (1,2)	1 (1,1.5)	*0.374*

[1] Scale: 0 (Never Important) to 2 (Always Important).

Table 36: Importance given by medical doctors to different types of data according to institution type.

	TYPE OF INSTITUTION		
	Hospital	Primary Care	
TYPES OF DATA[1]	Med. (p25; p75)	Med. (p25; p75)	*p*
Administrative Data	2 (1,2)	2 (1,2)	*0.787*
Alerts (allergies. etc.)	2 (2,2)	2 (2,2)	*0.86*
Diagnosis	2 (2,2)	2 (1,2)	*0.415*
Treatments	1.5 (1.5,2)	1.25 (1,2)	*0.131*
Clinical Events	2 (2,2)	2 (1,2)	*0.338*
Lab Exams	2 (1,2)	2 (2,2)	*0.719*
Image Exams	2 (1.25,2)	1 (1,2)	***0.01***
Image Exams Reports	2 (1.75,2)	2 (2,2)	*0.306*
Other Exams	2 (2,2)	2 (1,2)	*0.215*
Patient Mng. Doc.	1 (1,2)	2 (2,2)	***0.001***

[1] Scale: 0 (Never Important) to 2 (Always Important).

report rather than to the exam itself when compared with Emergency MDs. This attains statistical significance in Angiography (p=0.011), CT (p=0.006) and Echography (p=0.006). MRI and Osteodensitometry follow this tendency although in a lower degree (Table 37).

Also, it is observed, with significant difference, that the PCMD give more importance to reports than to the image itself for X-ray (p=0.046) , Angiography (p=0.008) , CT (p=0.001) , Echography (p=0.001), MRI (p=0.019) and Osteodensitometry (p=0.032) (Table 38).

Table 37: Relative importance between Imagiology exams images and reports among different profiles of doctors.

	PROFILES		
CASES WHERE	Outpatient	Emergency	
REPORT IS MORE IMPORTANT	n (%)	n (%)	*p*
X Ray	5 (16)	0 (0)	*0.053*
Angiography	11 (34)	2 (7)	***0.011***
CT	10 (31)	1 (3)	***0.006***
Echography	10 (31)	1 (3)	***0.006***
MRI	9 (28)	3 (10)	*0.108*
Osteodensitometry	10 (31)	5 (18)	*0.24*

Table 38: Relative importance between Imagiology exams images and reports among different institutions types.

	INSTITUTIONS		
CASES WHERE	Hospital	Primary Care	
REPORT IS MORE IMPORTANT	n (%)	n (%)	*p*
X Ray	1 (2)	4 (18)	***0.046***
Angiography	4 (17)	9 (41)	***0.008***
CT	2 (5)	9 (41)	***0.001***
Echography	2 (5)	9 (41)	***0.001***
MRI	4 (10)	8 (36)	***0.019***
Osteodensitometry	6 (15)	9 (41)	***0.032***

5.3.4 *Discussion*

Although MDs refer that in Hospitals the access to a broad type of patient data is a reality, this is not so present in the Primary Care Institutions. The barriers regarding access to external information are still in place and the fact that a reduced amount of MDs tries to access information outside their institution along with "old" communication means used to obtain it, reveals that there is still a long way to travel for making information available.

This reality clashes with the consensual view that the advantages for having access to external information are important and may have significant impact in the daily practice. However, having a deluge of information simply dumped into the MD desktop may have a counterproductive effect, a fact that may be inferred from the overall opinion that having access to external information may also reduce the time available for decision making. This evidence accompany the one presented in Shapiro et al. study [184] where it is also stated the opinion that time for decision-making is reduced when more information is available. In this sense, a priority oriented collection and enhanced presentation of all data may help to reduce this notion.

From the results, it is possible to observe that HMD tend to give more importance to hospital sources while PCMD do to primary care sources, suggesting that MDs tend to prefer data originating from institutions at the same level of care.

Although there is no information types getting discarded along the multiples profiles of health professionals the importance of some of them is highlighted in one profile more than others (e.g., Alerts that are more important in Emergency and Outpatient than in Inpatient settings). Another example is Patient Managed Data that is more relevant to Outpatient profiles. This type of data is mostly produced at this level of care and for this level of care hence the relevance being more restricted.

In image exams we observe an relative higher importance given to it by Emergency MD suggesting the need for having some information available as soon as possible, this is particularly evident in x-rays where no Emergency profile MD revealed that the report is more important than the exam itself. On the other hand, Outpatient profiles MD may prefer to rely in a specialist report rather than in their interpretation of the exam itself. This evidence gets stronger when MDs are grouped by institution of origin where PCMD give, in percentage, a higher importance to the exam reports.

By looking into different MD profiles, we get insight into possible differences regarding information need and usage. Approaches to HIS modelling and interface design may benefit from these findings by taking in consideration the existence of differences in MD profiles regarding the same types of data that can be presented in a more personalised fashion also the variability regarding sources of information preferred suggests that integration efforts may be directed to certain sources that are given higher importance than others.

Some limitations of the study rise from: the number of MDs in the Inpatient profile being low which may prevent a stronger statement regarding profiles heterogeneity; the length of the questionnaire that may have discouraged some MDs to complete the survey; and the study sampling being selected from institutions geographically near to the second city of Portugal, if answers from further inland institutions where included additional idiosyncrasies could rise as less health resources are available (e.g., Complementary means of diagnose).

5.4 SUMMARY

From these studies it was possible to highlight a relevant mobility pattern in the population and in the use of geographically distributed healthcare resources, this scenario may produce relevant financial impact if information is not shared among health institutions. It was also evidenced that when information sharing is at stake it is important to consider that for each scenario the information needs changes along with their time constraints and that these issues should be taken in consideration in the development of health system integration approaches. In the next chapter an extension to the previous intra-institutional health data agent based integration system is described aiming at extending the reach of the legacy system integration efforts outside the institutional barrier.

5.5 OUTCOMES

1. G. M. Bacelar-Silva, **P. M. Vieira-Marques**, and R. J. Cruz-Correia. Identifying relevant data along selected medical care situation. 6th Iberian Conference on Information Systems and Technologies (CISTI), Vol I, pages 350–353, 2011.[6]

6 AGENTS BASED INTER-INSTITUTION DATA INTEGRATION

In the previous chapter it was described that the population reveals relevant migratory movements with temporary or permanent changes in residency, and that patients may resort to several health institutions during their lifetime. Most of these health institutions are located within the boundary of the patient's usual "habitat" while others may lay scattered around the country or even abroad. In this scenario enhancing information availability is important and efforts should be made in order to make existing data delivered to Health Professionals (HP) at the point of care. However, it was also identified that different profiles of Medical Doctors tend to attribute a variable importance to different types of data, in different contexts of care, hence care must be taken to provide ways of prioritising and presenting data. In the following chapter, it will be described a system aiming at extending the previous efforts of health data integration within a single health institution to a wider reach outside the institutional border.

6.1 SCENARIO

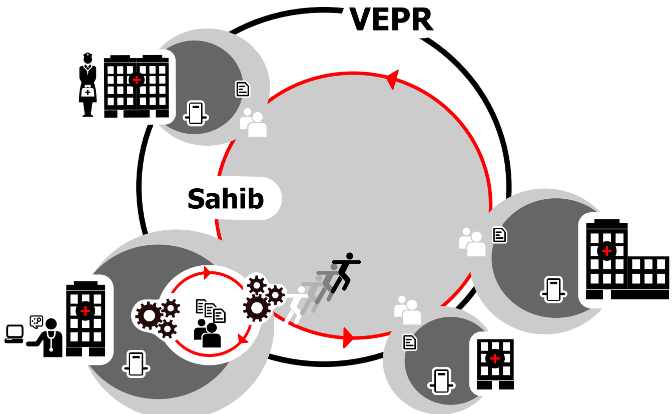


Figure 25: SAHIB - an inter-Institutional VEPR.

Pushing forward the previous efforts of making available at the point of care, a more comprehensive set of clinical information, including not only the local institution generated data, but also complementary data proceeding from other institutions that might have been visited by the patient. Thus, the reach of the previous system MAID is extended beyond the institutional border, moving towards an inter-institutional data integration model that provides a wider geographical reach VEPR.

In this context, SAHIB (Secure Agent based Health Information systems Brokering) aims to address a scenario that include health institutions of different types and sizes (e.g. Hospitals and Health Centres), located throughout the country, taking in consideration that several HIS may exist locally, that they might be implemented using different technologies and managing multiple types of data. It intends to bridge the health institutions information islands by providing an agent based technological approach that may overcome some of the integration issues (e.g. data accessibility, systems unavailability). The scenario addressed is based on the presumption of an intra-institution health information management system existence (MAID or other).

The approach proposed is driven by clinical events, in the sense that information retrieval actions will be triggered by scheduled events like outpatient and inpatient events and admissions through emergency.

6.1.1 *Actors*

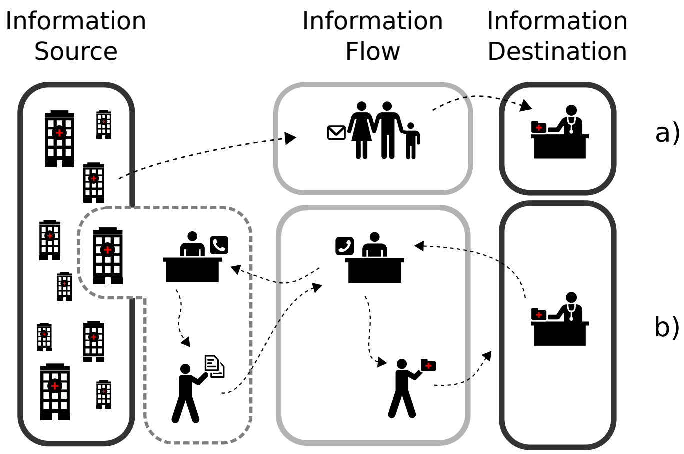


Figure 26: Scope, main actors and information flow.

The main actors in this system are the HP and the administrative personnel from health institutions that are responsible for gathering remote information and making it available at the point of care.

Indirectly, patients are also addressed as, usually, they are the main carriers of external information and in some cases the only mean for accessing other health institutions data. It can also be included in the set of actors the remote systems that usually do not have an access point for data exchange.

In Figure 26, it is depicted the usual process of access to external information. This process, usually, only takes place regarding previously scheduled events and usually after the first consultation when the patient informs the HP that, for instance, he was subjected to a surgery procedure or made exams in another institution. Scenario a) is the more common scenario were the patients are the ones carrying there information, however usually limited to lab exams or referral letters from other doctors. It is not usual for the patient to have access (although they are they legal owners) to other type of more clinical documents. Scanario b) takes place when a doctor ask for information from another health institution, it usually involves the archive personnel or some other administrative department from both health institutions to articulate efforts in order to make available the information to the doctor.

6.1.2 *Environment*

The system's environment is highly distributed one, with many institutional islands of information laying in independently developed systems using multiple technology and information models.

In most of the cases the institutions involved will be within a private network infrastructure managed by the public administration, however, in other cases like private hospitals and laboratories, or even in particular scenarios that go beyond the national borders, they may only be accessible through a more public network infrastructure.

Admission events (e.g. emergency, inpatient, outpatient) will drive the system workflow hence an articulation with the institutional administrative system is necessary.

Patient data collected should be accessible by HP through an integrated institutional web interface and will provide besides the clinical data, information regarding data provenance.

6.1.3 *Data Integration Model*

As institutional systems may vary regarding data structure and technologic infrastructure, data may be collected following two approaches: Document Oriented and Data Oriented making use of openEHR standard support.

Regarding patient identification the national patient identification number (num_utente) is used as it is the number that is more transversally available throughout the institutions. Data is made available until the scheduled event and then is archived for auditing purposes only. If and when data is again needed, it should be recollected enabling the system to cope with the advent of changes occurring since the last collection actions.

6.1.3.1 *Document Based*

In this model, like in MAID, the integration is achieved at the document level, each query to the remote system will generate a snapshot of the available data produced within a given time frame (e.g. the last six months). The document is associated to a specific department and classification is done around major groups of documents.

6.1.3.2 *Data Based using openEHR querying language*

In this mode the integration is based on data itself and is supported by openEHR constructs. The data exchange between openEHR based systems is based on performing AQL queries against the remote repository and retrieving openEHR extracts based on a common set of foundation archetypes. This approach will enhance interoperability and may accommodate the results from the previous chapter regarding the variable interests by HP in certain types of clinical information by having "personalised queries" according to a type of event or eventually according to a type of HP.

Collected data will embody the archetype structure (in the most basic approach an EHR_Extract element), the system that collects data may represent them directly without concerning about the contents strutter and leaving to the HP the interpretation of the provided data. However, being structured data, it may be subject to further usage like cross population studies.

6.2 SYSTEM MODEL

SAHIB's main actors and their actions (Figure 27) are modelled as autonomous agents that share a common goal of providing a more comprehensive VEPR.

This model builds upon the previous work on intra-institutional data integration system (MAID) described in Chapter 4, extending it with new features originating from the higher complexity of the new scenario addressed (e.g. institutional systems heterogeneity, network and systems availability). In continues to explore the characteristics of the multi-agent systems paradigm adding the mobility to the set of basic properties of agents in order to address network unavailability or time consuming data queries. The model includes internal agents

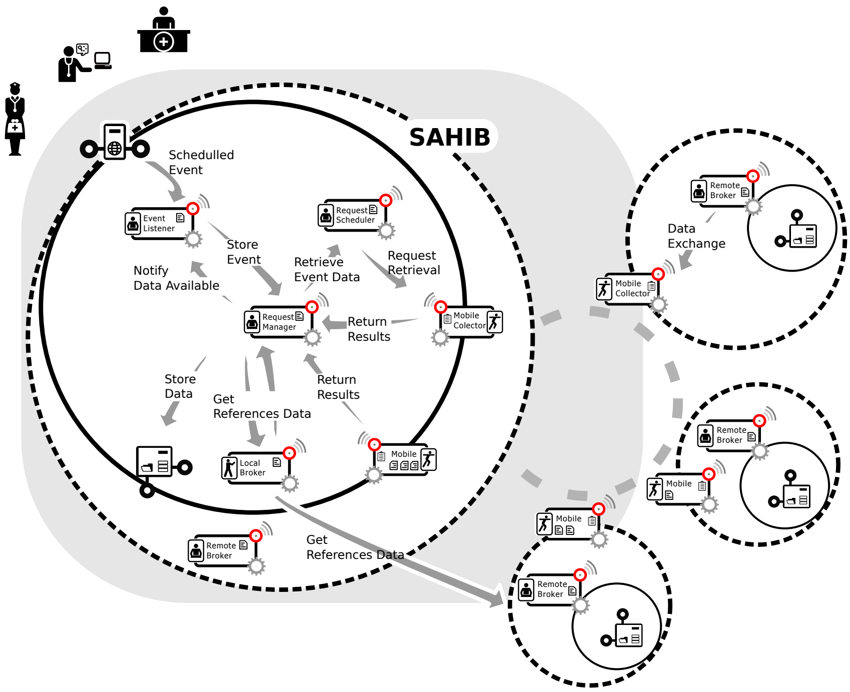


Figure 27: Agents community detail.

for local process management and external agents responsible for the interaction with incoming agents and their requests. All agents act on an asynchronous and autonomous way.

The system model is divided into three main processes :

event scheduling: This process will manage registered events and their scheduling constraints.

data discovery: With registered events are scheduled for execution, the discovery of new data will take place by having agents visiting several health institutions and querying for new patient data.

data retrieval: New patient data found will be copied over to the local system and made available to the requesting users.

Each of the described processes is decomposed and mapped into a set of agents. The articulation of these agents and their actions reproduce the necessary steps that must be undertaken to accomplish the purpose of finding and collecting new data in a timely fashion.

6.2.1 *Event Scheduling Process*

As administrative personnel register new events, they are transmitted over to the systems and pre-processed for collecting actions' scheduling. In this phase three types of agents are involved: Event Listener Agent (ELA); Request Manager Agent (RMA) and Request Scheduler Agent (RSA), they are described in the following subsections.

6.2.1.1 *Event Listener Agent*

The ELA acts as the interface for the institutional patient management system exposing, to the local HIS, an event registration service.

The institutional patient management system sends a message to the ELA informing that a new clinical event was scheduled and provides for all the relevant details regarding patient identification and other relevant event data (e.g. type and date of event).

The ELA will return to the requesting system an acknowledgement that the request will be processed and a request id that can be used for eventual updates on event data (e.g., cancelation, postponing).

In the next stage it will then provide the received data to the RMA for further processing. The main ELA interactions are described in Figure 28.

The message exchange between the HIS and the scheduling service is structured as set of fields relevant for processing and managing the request.

It includes a section for Patient Identification, with an element for the number and another for the type of identification (e.g. National Patient Identification Number), a section for identifying the HP that will consume this data including the HP number,the type of number (e.g. Medical doctor professional number) and the type of HP (e.g. Medical doctor, nurse), the purpose of the request (e.g. Clinical Events, Insurance) and the time constraints of the request (Listing 3).

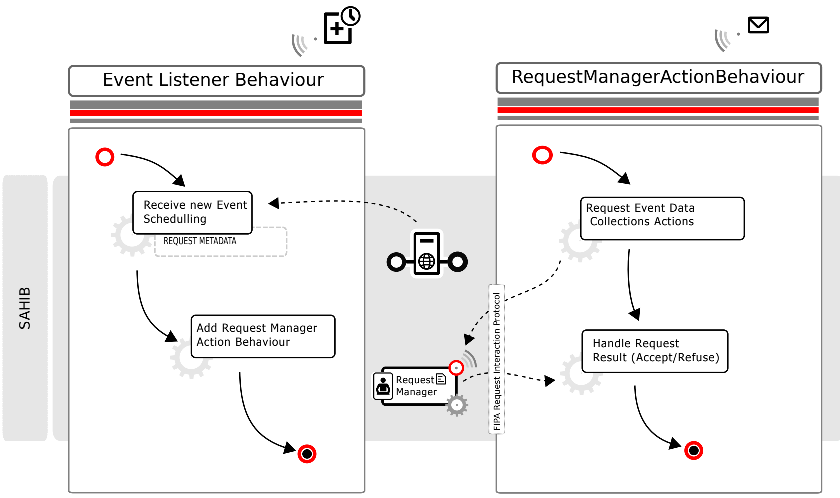


Figure 28: Event Listener Agent main actions and interactions.

Listing 3: Scheduled Event Metadata.

```
<requestMetaData>
  <requestAction>GetData</requestAction>
  <patientID>987654</patientID>
  <patientIDType>1</patientIDType>
  <userID>123456</userID>
  <userIDType>1</userIDType>
  <userRole>1</userRole>
  <requestPurpose>Patient Care — Outpatient</requestPurpose>
  <dateDataNeeded>2013/10/10 10:10</dateDataNeeded>
  <dateDataMaxUsage>2013/10/10 11:11</dateDataMaxUsage>
  <queries>
    <queryList>
      <query type="aql">
        <name>Blood Pressure Measurements</name>
        <code>SELECT c as comp FROM Ehr e[ehr_id/value=\$current_ehr_uid] CONTAINS VERSION v
              CONTAINS COMPOSITION c[openEHR-EHR-COMPOSITION.encounter.v1]</code>
      </query>
      <query type="aql">
        <name>Abnormal Blood Pressure Measurements</name>
        <code>SELECT
          c/name/value as composition_title,
          obs/data/origin/value as measurement_time,
          obs/data/events/data/items[at0004]/value as systolic,
          obs/data/events/data/items[at0005]/value as diastolic
          FROM
          Ehr e[ehr_id/value=$current_ehr_uid]
          CONTAINS COMPOSITION c[openEHR-EHR-COMPOSITION.encounter.v1]
          CONTAINS OBSERVATION obs[openEHR-EHR-OBSERVATION.blood_pressure.v1]
          WHERE obs/data/events/data/items[at0004]/value/magnitude > $limit
          ORDER BY c/context/start_time/value</code>
      </query>
      <query type="doc">
        <name>Available Data</name>
        <datebegin>2008/06/01</datebegin>
        <dateend>2014/06/01</dateend>
      </query>
    </queryList>
  </queries>
</requestMetaData>
```

There is also a section for including the queries that are relevant for the event type where the data will be used, these can be of two types: doc or openEHR AQL. It is possible to specify a date interval for the queries that in case of AQL based queries is embedded in query itself.

With the reception of new events messages, an interaction with the Request Manager agent will occur following a FIPA Request Interaction Protocol and messages are exchanged containing the data provided.

6.2.1.2 *Request Manager Agent*

The RMA plays a central role in the system as it is in charge of triggering the other system's processes (Data Discovery and Data Retrieval).

In the Event Scheduling process, this agent is in charge of receiving the requests from the ELA and storing the requests' data in the database for future scheduling actions.

As the request metadata is stored in the database, the RSA is notified about the existence of a new request.

Simultaneously the RMA is also in charge of receiving the collected data references, storing the results references and notifying the Local Broker Agent (LBA) that new references are available for retrieval.

The main RMA actions and interactions are described in Figure 29.

The RMA message exchange follow a FIPA Request Interaction Protocol with ELA, RSA and the LBA.

The message exchanged with the ELA during the Event Scheduling phase includes an object containing the metada from the request.

The message sent to the RSA during the Event Scheduling phase includes an object containing the metada from the request.

The message exchanged with the MCA during the Data discovery includes an object containing the collected references along with information regarding the location and the date of collection.

The message exchanged with the LBA includes the request id for identification of references to be retrieved.

6.2.1.3 *Request Scheduler Agent*

The Request Scheduler Agent (RSA) is in charge of periodically (e.g. at the end of the day) checking the database for upcoming events. In order to address the urgent requests like the ones originating from emergency events where the date to be made available is usually short, the agent's actions will also be triggered by RMA messages so that the necessary actions can be pursued immediately, otherwise it waits until the threshold (e.g. one week before) for the event is reached.

When the threshold is reached it calls for the Mobile Collector Agent (MCA) which will be responsible for data search and retrieval actions. It will provide the MCA with a request ID, the patient ID (e.g. a national patient identification number), the time constraints of the event, the queries requested, a set of locations ordered by geographical proximity to be visited and and a set of security credentials.

The RSA constructs two itineraries, one set contains the remote systems that are known to have patient information and other set with the other institutions retrieved from the directory service agent (ordered by size and relevance according to the origin of the patient and the type of event). These itineraries are delivered to two independent MCAs, one will have only the first set providing

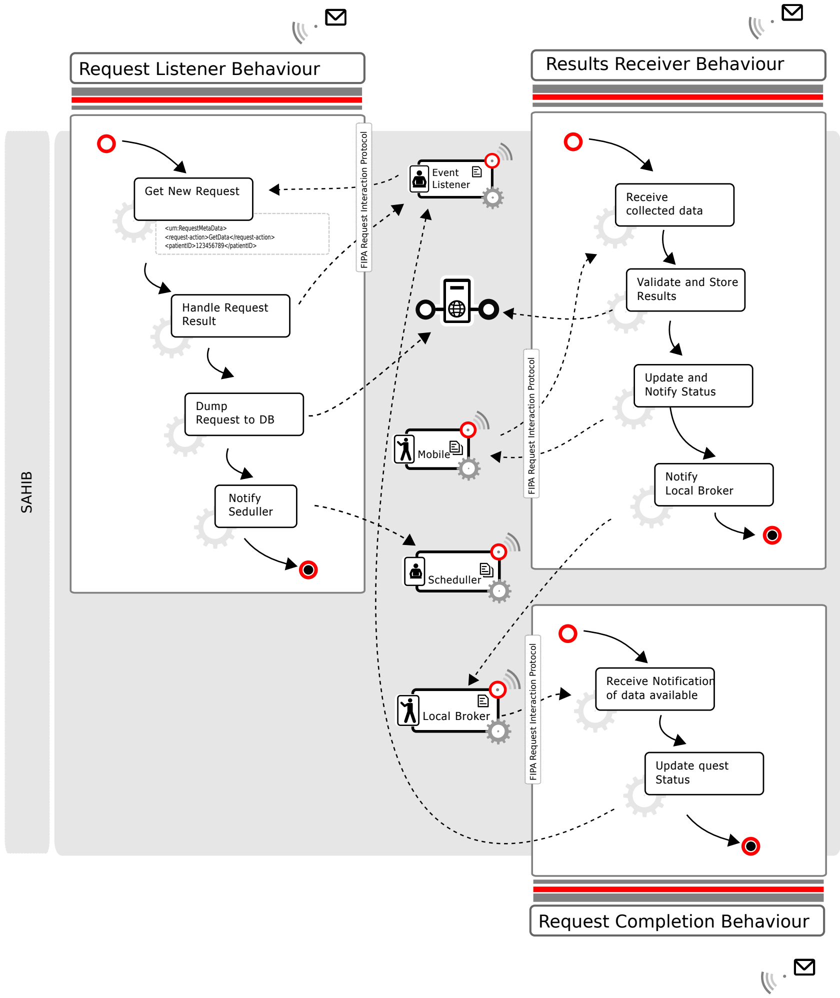


Figure 29: Request Manager Agent main actions and interactions.

a more directed search, the other MCA will have both sets in order to avoid visiting the same places as the first MCAs if for instance a known location is provided by a remote platform as a new location.

The main RSA actions and interactions are described in Figure 30.

The RSA receives messages from the RMA that include the new scheduled event request identification.

6.2.2 *Data Discovery Process*

As events are scheduled data search actions are triggered and several remote institutions will be visited in order to find additional patient data.

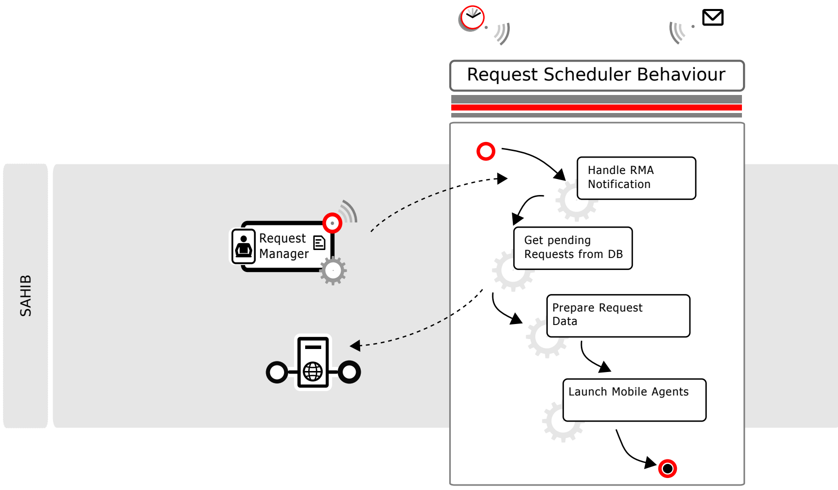


Figure 30: Request Scheduler Agent main actions and interactions.

Two types of agents are involved (Mobile Collector Agent and Remote Broker Agent), they are described in the following subsections.

6.2.2.1 *Mobile Collector Agent*

The MCA is in charge of searching for information around health institutions and bringing it to the home system. The MCA is provided with a patient identification and a pre-determined itinerary that is followed sequentially and that can be enlarged by interaction with remote systems.

The Mobile Collector Agent (MCA) actions are guided by a finite state machine which include states for each of the stages needed to move around and to interact with remote and local agents (Figure 31). When the MCA reaches a new location, it will look for that institution Remote Broker Agent (RBA).

If the MCA can find the RBA it will then request access by providing a set of credentials and information, otherwise it will identify the next location and start the migration procedures.

According to the type of interface available (document based or data based) it will send the queries for information about the patient and if any available, it will receive references for data to be later retrieved.

After retrieving the data it will proceed to ask for any other locations known to the local system that might be complementary to the ones he owns from origin. The MCA will include the new locations at the end of his original itinerary after evaluating if they are already know by the home location. This list will grow has new places with information are discovered.

As the actions of data collection finishes, the MCA will evaluate it's next jump and repeat these stages for the next location.

Time constraints that are added to the original request influence the MCA decision to either continue the search and collection actions or to return to the home institution with the collected items.

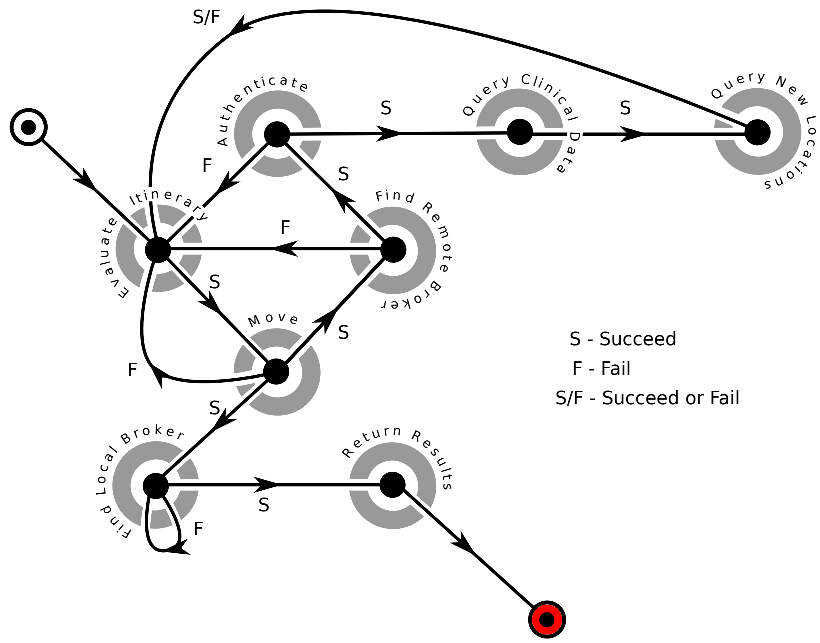


Figure 31: Mobile Agent Finite State Machine.

The procedure repeats until the MCA has visited all the locations on its itinerary or the time constrains have been reached.

In case of migration impossibility to a new location the agent will add this location to a list of locations that were not possible to visit and continues to the next location in queue. After the inicial set of locations is visited the list of platforms that evidenced problems is retrieved and added again to the itinerary. The agent will continue to try to visit these locations until success is attained or the threshold for return is reached.

When reaching the home institution it will find the RMA and returns all the data gathered for further actions of data references retrieval and the additional locations containing information for the that eventually were found along the way.

The Local and Remote interactions of the MCAs are described in Figure 32 and Figure 33.

6.2.2.2 *Remote Broker*

This agent resides in each institutional platform and acts as an interface agent to incoming agents looking for patient information Figure 34. It has to have a local HIS interface for data collection. And depending on the local HIS it will accept document references query or openEHR AQL queries.

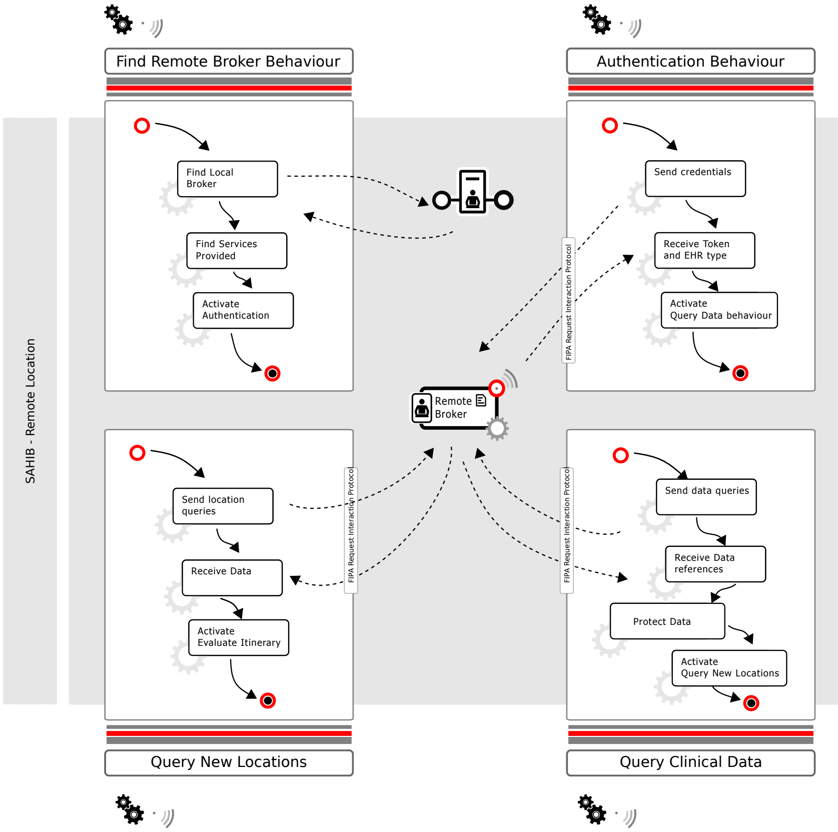


Figure 32: Mobile Collector Agent main actions and interactions at remote institution.

Listing 4: Request results data.

```
<reports>
  <report>
    <department>Clinical Patology</department>
    <reporttype>Lab Report</reporttype>
    <reportdate>02/10/2008 16:18</reportdate>
    <docref>987654321.pdf</docref>
  </report>
  ...
</reports>
```

Interactions with the MCA follow the FIPA Request Interaction Protocol and upon validating the authentication credentials it will provide the incoming agent with a token to be used in the next interactions and the type of available HIS (doc or openEHR based). It will receive a message containing the request metadata and process the queries in order to make available the data references. It will store the data validity (time of consumption) in order to enable a background cleaning process to eliminate the processed data from the differed repository.

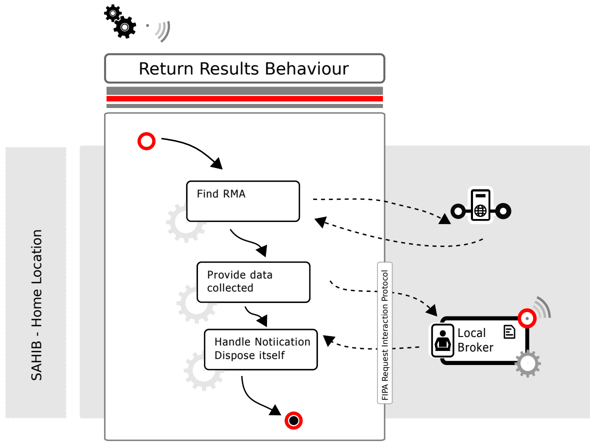


Figure 33: Mobile Collector Agent main actions and interactions at local institution.

6.2.3 *Data Retrieval Process*

In this process the data references collected by the MCA will be processed and information will be made available to the user.

For this purpose, one additional type of agent is involved (Local Broker Agent), it is described in the following subsections.

6.2.3.1 *Local Broker*

This agent is in charge of processing the results of the references retrieval process Figure 35.

After receiving a notification from the RMA with the request id, the LBA will use the stored references to retrieve the remote data through the remote interface web service. It will process the information according to the type of data retrieved incorporating it in the local HIS external data database and file system.

After the documents are retrieved a notification is sent to the RMA stating that the process is complete and the request status can be changed accordingly. After these stage they are made available to the end user (Doctor, Nurse, etc) through existing user interfaces.

6.3 SECURITY

Having agents that can move along several platforms, carrying sensitive data from one execution environment to the next, and acting on behalf of others, raise security requirements that must be considered. The authentication of users and agents is achieved by means of external identity-based PKIs and roles are provided by the local institution. Protecting in transit agent's code and content is achieved by adopting a model of self-protected mobile agents where agents carry their own protection mechanisms. The initial self-protected scheme allows

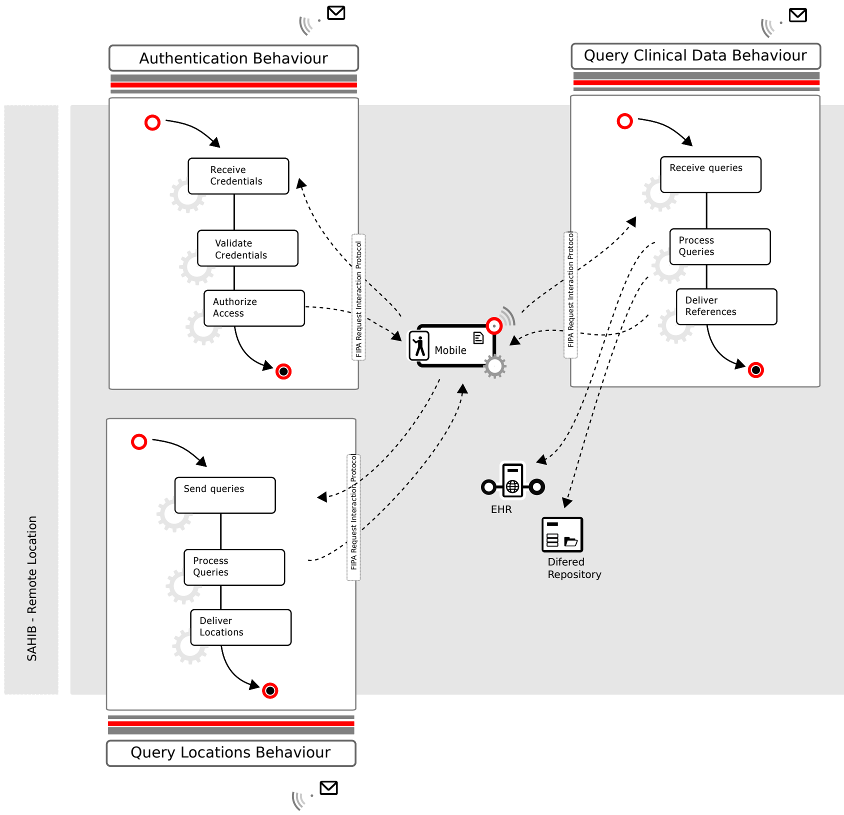


Figure 34: Remote Broker Agent main actions and interactions.

an agent with a static itinerary to implement its own security mechanisms to protect the code to be executed at every platform. A more detailed and complete description of the self-protection mechanism can be found in [185]. In this case the system makes use of semi-static itineraries, because at the beginning only a subset of the platforms the agent will visit is known (some of the platforms will be discovered during the journey of the agent). However, the original self-protection scheme does not support non-static itineraries. Because of this, it is adopted an adapted version of the original self-protection mechanism. Instead of encrypting the agent code only at the very beginning, it will be reciphered on every platform the agent migrates in such a way that only the next platform to visit will be able to access the code. This allows the agent to add new platforms dynamically into its itinerary.

Another important asset to protect is the information carried by the agent. It is important to shield agent's code but if the results carried by the agents are not protected against modifications or eavesdroppers the whole security is jeopardised. In our proposal, we adopt a scheme based on hash chains to protect agent's results. Similar mechanisms have been used before to protect agent data, as described in [186]. This type of protection prevents the results from being undisclosed or changed by unauthorised parties. The results on a platform are

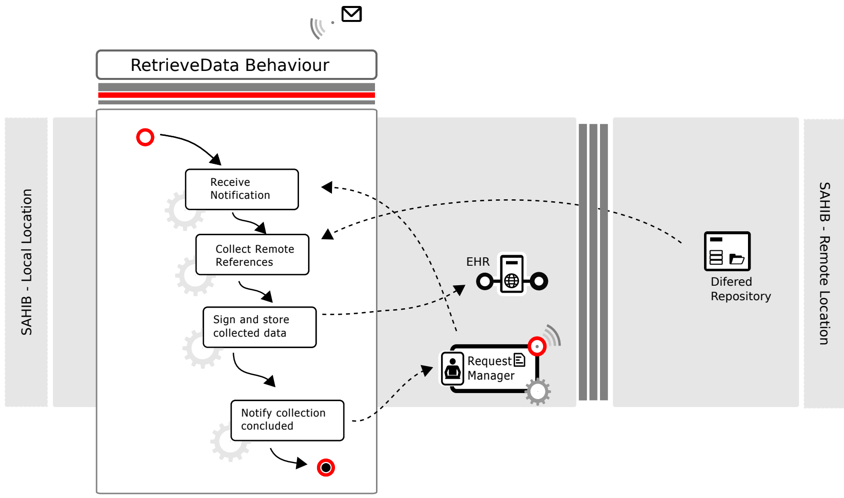


Figure 35: Local Broker Agent main actions and interactions.

encrypted using the public key of the owner thus protecting it until the final destination and additionally hash chains are generated and added to prevent replacement or removals.Only if malicious platforms colluded would it be possible to partially break the hash chain, thus permitting some alteration of the results.

The access control mechanism is an improvement of the basic RBAC in the MAID system. The RBAC subsystem proposed is composed of two main modules: the authentication module, and the RBAC module (role enabling and local role authority). While the RBAC module provides enforcement and management of RBAC policies.

Each domain or institution has a role authority, which is responsible for managing the local and global RBAC policies, and an enforcement point, called role enabling module. The role enabling module determines the roles of the user. The authority provides role information to the enabling module. The authority is responsible of the roles defined and managed in the institution so it may be queried by third parties, or authorities for other institutions about the roles it defines. These roles reflect the internal organisation of the institution, for example physician, nurse, department of radiology, etc. Local RBAC policies are enforced when the user makes the patient record request.

Besides local role enforcement there is a global role enforcement. Global role enforcement is produced, when a mobile agent arrives to a given institution. The agent has the role inherited from its user along with the purpose of the request, provided by the local institution, which is included in the RBAC policies.

The RBAC policy is local to the institution so a given institution could define its own format. Nevertheless the use of eXtensible Access Control Markup Language (XACML) language is a well known standard for access control policies, what ensures that the system will easily interoperate with other access control policies if needed in the future. Authentication may be based on a PKI or not depending on the specific environment. For instance, by health government department Certification Authority (CA), local hospital CA, SSO module, etc. This makes the authentication very flexible, an institution could use simple password based authentication, while another one could use more sophisticated methods of authentication.

6.4 USER INTERFACE

The previously described VIZ interface was extend to accommodate the necessary interfaces for enabling the event scheduling and collected information visualisation.

The data collected is shown aggregated by Health Institution and according to the type of document retrieval mode (depending on the type of remote HIS) (Figure 36). In the case of Document based mode they are shown as a list of records to be viewed individually and in the Data based mode associated with the queries requested. If the documents are openEHR EHR extracts the XML files are transformed into HTML through a conversion mechanism. As the collected data follows the openEHR reference model it may represent it directly without concerns about its contents, leaving to the HP the interpretation of the relevant information(Figure 37).

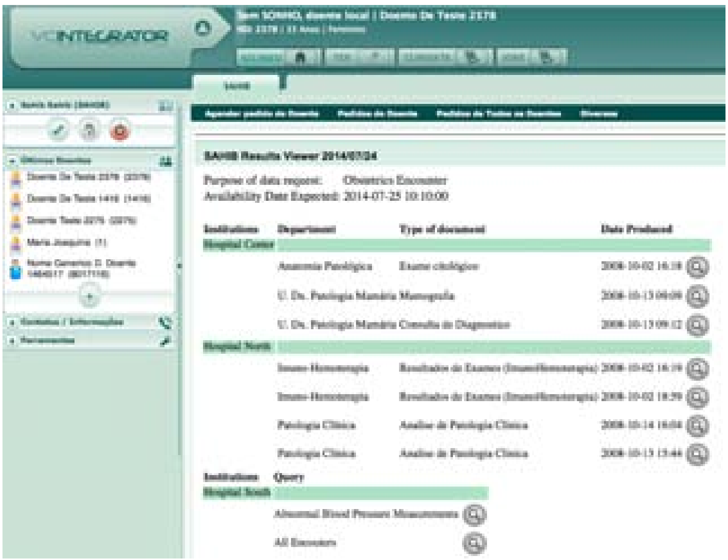


Figure 36: Detailed list of data collected for a particular patient.

6.5 IMPLEMENTATION

The model presented was implemented using the JADE (vs 4.2) framework and its API, agent communication and mobility services, are based on IEEE-FIPA standards and Interaction Protocols. Agents mobility is provided by JIPMS add-on which is based on sending Mobile Agents over an ACL Message that has

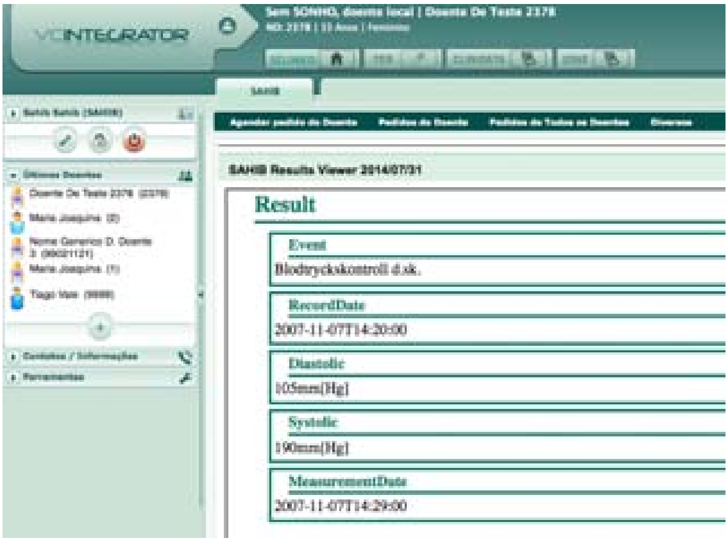


Figure 37: Example of an openEHR extract composition - Abnormal Blood Pressure measurements.

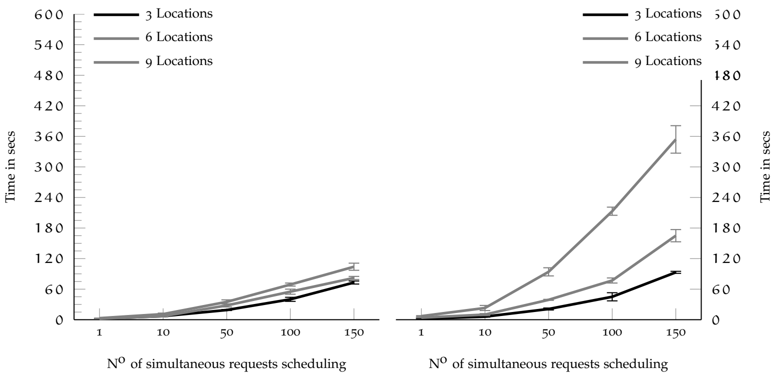
been integrated with the Intra-Platform Mobility Service in Jade. This service allows an agent to migrate between containers and platforms by using already existing methods in a standard way (through IEEE-FIPA compliant mechanism). This standardised mobility facilities will be used for clinical information gatherer agents to travel over the multiple existing systems. For web services implementation the Apache Tomcat server ([187] and JADE WSIG add-on was used for web services integration.

For data persistency, an Oracle 10i server was used for storing and retrieving the necessary data (patient, events and document management). The openEHR simulated EHR repository is implemented using a XML database (eXist-db [188]) based on the Linkoping University Educational EHR Environment (LiU EEE) [189].

The system was deployed in four virtualised servers and the the purpose of creating a simulation scenario one of the servers was used as the requesting institution and the rest of the set as potential data providers. The main server was a double processor Xeon 2.40GHz 2GB of memory, two servers with single processor Xeon 2.40GHz, 500Mb of memory and the third a single processor Xeon 2.0GHz with 500Mb of memory.

For simulating the scheduling of events it was used a properly anonymised set of encounter data from an obstetrics system deployed in a major Portuguese Hospital. A cron would simulate the scheduling of the reports based on recorded time of the event. The events were aggregate for each hour and scheduled through the systems's scheduling service. The simulated remote institutions provide simulated reports for the virtual patients. The system was left to run for a week the simulated events reproduced the pattern of events scheduled during a each hour of each day. The expected number of scheduled events and retrieved reports references was reached.

For additional testing regarding time to find and retrieve document references to the home system was tested for two scenarios: no results returned and a fixed set of one thousand results references per location amounting for a total of 60kb (these are compressed by the mobile agent giving an average of 7kb per location visited).



(a) 0 references from each location (b) 1000 references from each locations

Figure 38: Data search and references retrieval time measurements

For each of those scenarios the system was tested with three, six and nine locations. As only three servers were available the agents would repeat the visits for each server. Time since request scheduling and request complete was measured and recorded.

Each test was run five times, average and standard deviation of time differences between database records of date of request scheduling and date of request references storage were calculated. The results obtained are presented in Figure 38.

The increase in the number of agents, the number of locations to visit and the consequent message size leads to an increasing processing time in each location thus increasing the overall report availability (approximately 2 times for each 50 sequential requests) however still in a reasonable time frame. The amount of references would vary from institutions and it would be unlikely for a patient to have one thousand reports in each institution. Nevertheless a more realistic environment (additional standalone servers with more processing power) and an increased number of hourly requests would provide better assessment of the implementation. The increasing time also originates from having repeated locations visited by the agents which along with the low servers resources available in the provider servers may produce increased contention in the migration process.

6.6 CONCLUSIONS

As seen before, clinical data is prone to be highly distributed among several institutions and the knowledge of its existence and their actual location is often unknown.

In this work it is described a system that aims to extend the reach of the previous intra-institutional VEPR system described in Chapter 4 in order to enhance data integration coverage.

Like MAID it intends to optimize the health institutions information flow by providing a secure agent based clinical data discovery and transport mechanisms allowing each institution to have access to external complementary patient clinical information addressing issues like inter-institution legacy patient health data integration and unavailable on-line remote data retrieval.

Agents provided a seamlessly approach for mapping the main interveners in the process of getting external information and, by acting on behalf of health professionals, to anticipate the availability of clinical information.

Based on the distributed knowledge that each institution has, the system provides a way of finding out new sources of patient information without relying on the patient's volatile memory.

The document based data integration model addresses legacy systems interfacing issues through requesting a snapshot of the data available following the local format. This approach provides a pragmatic and straightforward way of increasing the information available without the need to map every type of data, data structures and semantics proceeding from different proprietary systems.

However, embodying the need for standards awareness, it also provides a mechanism for interfacing openEHR based systems. By using openEHR AQL and the fact that it expresses the queries at the archetype level(i.e. semantic level) it explores the stability, robustness and flexibility of archetypes the standard constructs for openEHR.

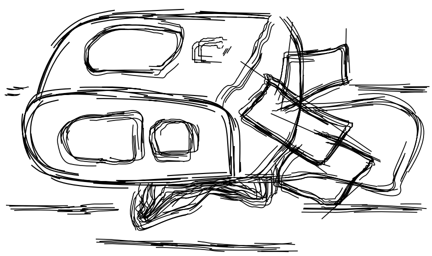
This approach allows the pursuit of an approach for semantic interoperability through semantic queries that enable to retrieve standardised data independently of the remote system underlying data model and persistence approach. The queries carried by agents can be created according to health professionals or event specific data needs providing a more personalised data collection.

6.7 OUTCOMES

1. **P. M. Vieira-Marques**, R. J. Cruz-Correia, S. Robles, J. Cucurull, G. Navarro, and R. Marti. Secure integration of distributed medical data using mobile agents. Ieee Intelligent Systems, 21(6):47–54, 2006.[3]
2. **P. Vieira-Marques**, A. Cunha, L. Antunes, R. Cruz-Correia, and A. Costa-Pereira. A first approach for a regional wide VEPR. Healthinf 2008: Proceedings of the First International Conference on Health Informatics, Vol 1, pages 215–218, 2008.[4]
3. Silva-Ferreira, P. R.; Patriarca-Almeida, J. H.; **Vieira-Marques, P. M.** and Cruz-Correia, R. J. Improving expressiveness of agents using openEHR to retrieve multi-institutional health data Feeding local repositories through HL7 based providers Aisti-Assoc Iberica Sistemas e Tecnologias Informacao, 2012, 885-889 [8]
4. **Vieira-Marques, P.**; Patriarca-Almeida, J.; Frade, S.; Bacelar-Silva, G.; Robles, S. ; Cruz-Correia, R. OpenEHR aware multi agent system for inter-institutional health data integration Information , 2014 9th Iberian Conference on Systems and Technologies (CISTI), 2014[10]

Part III

UNPACKING



Log Entry #3 : The end of this journey is near, we have gathered some items along the way: a system that will provide enhanced access to information within a health institution; some postcards of how patients use health institutions and how different profiles of health professionals use patient information and an extension to the system that will enlarge the integration boundaries to a inter-institution scenario. In the last stages we will reason about where this journey has taken us and how it may have contributed to fulfil the initial expectations.

7 CONCLUSIONS AND FUTURE WORK

In previous chapters the work developed around the research questions of how agent based technology can enhance health information systems integration has been described and discussed. In the next sections a summary of the main achievements will be provided and some recommendations derived from the work developed will be described. Finally, are proposed paths that may be interesting to follow in the future.

7.1 CONCLUSIONS

The work presented in this thesis has as starting point the issues that surround the efforts of making available a more complete set of patient data at the point of care and the proposal of an agent based approach for addressing them.

In Chapter 2 it is made a review of the health care data organisation shapes evolution and the main difficulties and advantages of having a more wide access to patient data. Concepts of agent technology are described and its adequacy to address health care information problems is evidenced.

From the systematic review of agent based system in the literature described in Chapter 3 it was possible to characterise and define a set of usage profiles that reinforce the notion of agents being suitable for addressing issues in the healthcare domain. The problems addressed are diversified and the main interveners in the healthcare process are represented. The proposed systems cover several levels of geographical range and different types of health institutions. From the implementation point of view, some requirement gaps were identified, particularly the ones related to missing security and health information standards awareness. It was further evidenced that frameworks supporting the systems' prototypes development tend to diverge from a self developed agent infrastructure to available community developed frameworks where Jade framework is leading in the preferences of researchers.

From this point forward, the work focused in modelling a health data integration approach within a single health institution where a scenario of multiple departmental legacy systems coexisted and the need for a single point of information access was necessary (Chapter 4). The agent based approach provided the necessary constructs to address the information flow and requirements of providing an integrated view of available data throughout the hospital by aggregating into a single user interface several departmental information systems data. Agent technology and decomposition of the workflow allowed to model quite transparently the existing manpowered information flow creating the opportunity for resource reallocation and improvement in document availability.

By having the integration of data performed at the document level, it is possible to reduce the need for a costly and time-consuming operation of mapping all available data structures and semantics providing a straight forward way of making information available without interfering with the running departmental systems.

The split process of retrieving document references and documents provided for a flexible approach allowing to have a faster updated patient document index anticipating the information needs at the point of care.

Additionally, the approach of retrieving actual documents from remote system allows the creation of a historic repository of patient data that can be used for auditing purposes regarding the amount of information available at the time of decision making and also to be used in the advent of a departmental system ceases to exist.

The system has been deployed and is still in use within a major health institution. From the experience in the deployment, some issues came out regarding the use of agent technology that need to be taken into consideration: agent actions monitoring was not easy and, although the system has a low incidence of non-deterministic behaviour, this proved to be one difficulty to overcome.

Having real data accessible provided a valuable asset in order to test additional features to be added to the system and allowing the creation of a realistic simulation environment allowing for further testing of system changes and to access their reliability. The environment was used to test an alternative approach to the static behaviour by proposing an algorithm for the document reference lists agent actions scheduling based on past knowledge of reports production from each department. It was also proposed an approach that takes in consideration this past information of departmental production to make agent actions more adaptable to the dynamic reality, allowing the reduction of time to make data available without incurring in unnecessary stress of the departmental systems.

In Chapter 5, a series of studies were conducted in order to better describe the nature of potential and effective patients mobility patterns and to study if health information may be stratified according to their importance or relevance to different health professionals.

The analysis of mobility evidences the spread of information for patients. It has been identified that although the main locations to look for information are closer regarding geographical location a wider search scope should also be taken in consideration.

From the second and third study it has been possible to provide evidence that the types of data needed varies according to particular clinical contexts and health professionals. The health professionals opinion regarding access to external data is positively strong and considered to have an impact on their daily routine. However, having just access to a big set of data may not be enough and in some cases can be counter productive as different profiles of health professionals attribute variable importance to the same type of information in different contexts of care. This suggests that the profile of the user that is going to use the information is important and should be taken in consideration when developing health information systems that present patient data proceeding from different sources.

Having in mind the characterisation of patient mobility and the importance and value attributed to having extended access to external information by health professionals, Chapter 6 describes an extension of the intra-institutional health information integration system to a multi-institutional scenario. This extension follows the document based approach regarding data integration and makes uses of agents mobility to address heavy queries issues and network instability that may occur by proposing the necessary additional security measures raised by this approach.

It also proposes a standard based approach using the same agent based infrastructure already developed for the document based approach. Although dependent on source systems having an standards based health record implementation or a more complex translating mechanisms, the approach adopted provides a non standard based system to incorporate and make available data resulting from structured openEHR AQL queries. This provides the grounds for a more personalised approach on the types of data that should be gathered from remote systems.

In general, the developed systems evidenced that agent technology is able to cope with the requisites of integrating data within a healthcare institution and to be deployed in a real world scenario where legacy systems coexisted without. The integration model proved also to be scalable and extendable to a multi-institutional scenario and even to emergency scenarios [190]. Being able to act on behalf of health professionals, is a strong point in favour of using agent based systems. This has been further evidenced by the systems developed providing a way for anticipating the users need of information and making it available in a timely fashion.

7.1.1 *Main ndings*

The next items summarise the main findings resulting from this work. From the characterisation of the environment and the development of systems aiming to enhance information availability at the point of care it can be referred that:

- security and health standards awareness should be taken in higher consideration.
- Jade is becoming the more used framework.
- agent technology is adequate for addressing health information systems integration problems and has proven to be well adopted in a real world scenario and easily scalable to include new sources of departmental data.
- the agents actions may be guided to adapt to external systems behaviour reflecting the different profiles of each department work load. Through an adaptive algorithm, it has been possible to enhance an adaptive behaviour of agents actions by taking in consideration the past history of departmental document production.
- from a characterisation of the potential population usage of health resources it is reinforced the need for generalised information access and it is possible to identify search priorities that in a scenario where information is dispersed and there is no knowledge about its location, an information indexing method is necessary.
- health professionals consider that having a lot of data available is important but that will lead to a decrease in the time available for patients.
- data is not consumed indiscriminately and there are variations in the importance given to different types and sources of data by different types of

health professionals in different contexts of care. This evidences the need for a strong data oriented search approach and concept based presentation approaches to be made available to health professionals.

- from the extension of the intra-institution model for data integration to a inter-institutional system suggests the flexibility of agent technology to be scalable.
- is possible to protect roaming agents with a dynamic itinerary.
- in order for agents to embody more deeply their representative role of health professionals and reflecting the real world process of requesting external information the purpose for the request should be stated.
- by incorporating the openEHR awareness and exploring its query language it is possible to integrate simultaneously data from both legacy non standardised systems and openEHR standard based systems.

7.1.2 *Recommendations*

From an implementation point of view and given the experience while deploying an agent based system in a real world scenario some items are relevant to refer:

- Adopting security measures increases the confidence of health professionals to adhere to the system.
- Debugging and monitoring the system revealed to be one big challenge.
- The access to real data revealed to be important for the development of a realistic simulator where further tests can be performed and performance comparisons can be made between different implementation approaches.
- The real time based system testing implied long running experiments without the possibility of accelerating the system execution. This leads to a more time-consuming debugging operations. However, it is an important process particularly when exploring the autonomy of agents and there is the need to ensure a reliable and smooth system operation.

7.2 FUTURE WORK

Although the document based approach provides a rather straightforward way of increasing the information available, it is not enough to respond to the healthcare community needs of relevant data sets and configurable interfaces which must rely on structured data. Governments are becoming aware of the importance of having a well structured articulated and future proof health information systems and are actively promoting the standards adoption by imposing market policies and standards compliance. In this sense, it is important that agents can embody standardised data and constructs more deeply thus providing ways of making agents able to manipulate standardised data.

Although there are some examples of HL7 adoption not many experiences exist regarding the use of openEHR approach. The openEHR initiative provide a comprehensive approach aiming at providing the necessary constructs that lead to semantically interoperable health information systems. Having agents that can embody it's ontology and reason about sets of standardised data may provide new approaches for rich querying models which will enhance the usage of available health information not only for individualised care purposes but also for public health patterns identification and broader clinical research activities.

On the other hand, providing health professionals with ways of personalising their data integration queries (e.g. query building interfaces) strengthens their engagement in the construction of a more adaptable and useful heath information system. This approach may also provide an important extension of the system enabling patient access and promoting patient empowerment by providing alternative ways of data access.

BIBLIOGRAPHY

[1] R. Cruz-Correia, P. Vieira-Marques, P. Costa, A. Ferreira, E. Oliveira-Palhares, F. Araujo, and A. Costa-Pereira. Integration of hospital data using agent technologies - a case study. *Ai Communications*, 18(3):191–200, 2005.

[2] P. Vieira-Marques, R. Cruz-Correia, P. Costa, E. Palhares, A. Ferreira, and A Costa-Pereira. MAID: Multi Agent for the Integration of Data. *TÈkhne - Revista de Estudos PolitÈcnicos*, pages 73 – 85, 06 2006.

[3] P. M. Vieira-Marques, S. Robles, J. Cucurull, R. J. Cruz-Correia, G. Navarro, and R. Marti. Secure integration of distributed medical data using mobile agents. *Intelligent Systems, IEEE*, 21(6):47–54, 2006.

[4] P. Vieira-Marques, A. Cunha, L. Antunes, R. Cruz-Correia, and A. Costa-Pereira. A first approach for a regional wide vepr. *HEALTHINF 2008 - 1st International Conference on Health Informatics, Proceedings*, 1:215–218, 2008.

[5] J.H. Patriarca-Almeida, P.M. Vieira-Marques, and R.J. Cruz-Correia. Simulation environment for the optimization of the data retrieval capabilities of an agent based system in a healthcare setting. *Lecture Notes in Computer Science*, 6867 LNAI:124–132, 2011.

[6] G.M. Bacelar-Silva, P.M. Vieira-Marques, and R.J. Cruz-Correia. Identifying relevant data along selected medical care situation. In *Information Systems and Technologies (CISTI), 2011 6th Iberian Conference on*, pages 350–353, June 2011.

[7] J.H. Patriarca-Almeida, P.M. Vieira-Marques, and R.J. Cruz-Correia. Optimization of an agent based clinical data retrieval system. In *Information Systems and Technologies (CISTI), 2012 7th Iberian Conference on*, pages 1–6, June 2012.

[8] P.R. Silva-Ferreira, J.H. Patriarca-Almeida, P.M. Vieira-Marques, and R.J. Cruz-Correia. Improving expressiveness of agents using openehr to retrieve multi-institutional health data: Feeding local repositories through hl7 based providers. In *Information Systems and Technologies (CISTI), 2012 7th Iberian Conference on*, pages 1–5, June 2012.

[9] P. Vieira-Marques, R. J. Cruz-Correia, and Sergi Robles. Profiling roles and scenarios of agent application in healthcare - a systematic review. In *VIII Workshop on Agents Applied in Health Care*, 2012. held in conjunction with the 14th Conference on Artificial Intelligence in Medicine (AIME 2013).

[10] P. Vieira-Marques, J. Patriarca-Almeida, S. Frade, G. Bacelar-Silva, S. Robles, and R. Cruz-Correia. Openehr aware multi agent system for inter-institutional health data integration. In *Information Systems and Technologies (CISTI), 2014 9th Iberian Conference on*, pages 1–6, June 2014.

[11] R. Cruz-Correia, P. Vieira-Marques, A. Ferreira, E. Oliveira-Palhares, P. Costa, and A. Costa-Pereira. Monitoring the integration of hospital information systems: how it may ensure and improve the quality of data. *Medical and Care Compunetics 3*, 121:176–182, 2006.

[12] R. J. Cruz-Correia, P. M. Vieira-Marques, A. M. Ferreira, F. C. Almeida, J. C. Wyatt, and A. M. Costa-Pereira. Reviewing the integration of patient data: how systems are evolving in practice to meet patient needs. *Bmc Medical Informatics and Decision Making*, 7, 2007.

[13] V. Ferraz, G. Oliveira, P. Viera-Marques, and R. Cruz-Correia. Organs transplantation - how to improve the process? *Stud Health Technol Inform*, 169:300–4, 2011.

[14] Stanley J. Reiser. The clinical record in medicine part 1: Learning from cases. *Annals of Internal Medicine*, 114(10):902–907, 1991.

[15] Stanley J. Reiser. The clinical record in medicine part 2: Reforming content and purpose. *Annals of Internal Medicine*, 114(11):980–985, 1991.

[16] R. S. Dick, E. B. Steen, and D. E. Detmer, editors. *The Computer-Based Patient Record: An Essential Technology for Health Care, Revised Edition*. The National Academies Press., 1997.

[17] L. L. Weed. Medical records that guide and teach. *New England Journal of Medicine*, 278(11):593–&, 1968.

[18] R. E. Dayhoff and D. L. Maloney. Exchange of veterans affairs medical data using national and local networks. *Annals of the New York Academy of Sciences*, 670:50–66, 1992.

[19] Kristiina Hayrinen, Kaija Saranto, and Pirkko Nykanen. Definition, structure, content, use and impacts of electronic health records: A review of the research literature. *International Journal of Medical Informatics*, 77(5):291 – 304, 2008.

[20] ISO/TC-215. Health informatics - electronic health record - definition, scope, and context. published standard iso/tr 20514:2005. Technical report, International Organization for Standardization, 2005.

[21] D. W. Forslund, R. L. Phillips, D. G. Kilman, and J. L. Cook. Experiences with a distributed virtual patient record system. In *Proceedings : a conference of the American Medical Informatics Association AMIA Annual Fall Symposium. AMIA Fall Symposium*, pages 483–7, 1996.

[22] H. van der Linden, J. Talmon, H. Tange, G. Boers, and A. Hasman. An architecture for a virtual electronic health record. *Studies in health technology and informatics*, 90:220–5, 2002.

[23] F. Malamateniou and G. Vassilacopoulos. An implementation of a virtual patient record using web services. *Stud Health Technol Inform*, 95:158–63, 2003.

[24] P. C. Tang, J. S. Ash, D. W. Bates, J. M. Overhage, and D. Z. Sands. Personal health records: definitions, benefits, and strategies for overcoming barriers to adoption. *Journal of the American Medical Informatics Association : JAMIA*, 13(2):121–6, 2006.

[25] J. C. Wyatt. Hospital information management: the need for clinical leadership. *BMJ*, 311(6998):175–8, 1995.

[26] D. C. Kaelber and D. W. Bates. Health information exchange and patient safety. *Journal of biomedical informatics*, 40(6 Suppl):S40–5, 2007.

[27] R. Amarasingham, L. Plantinga, M. Diener-West, D. J. Gaskin, and N. R. Powe. Clinical information technologies and inpatient outcomes: a multiple hospital study. *Archives of internal medicine*, 169(2):108–14, 2009.

[28] C. P. Friedman, A. K. Wong, and D. Blumenthal. Achieving a nationwide learning health system. *Sci Transl Med*, 2(57):57cm29, 2010.

[29] M. E. Frisse and R. L. Holmes. Estimated financial savings associated with health information exchange and ambulatory care referral. *Journal of biomedical informatics*, 40(6 Suppl):S27–32, 2007.

[30] M. E. Frisse, K. B. Johnson, H. Nian, C. L. Davison, C. S. Gadd, K. M. Unertl, P. A. Turri, and Q. Chen. The financial impact of health information exchange on emergency department care. *Journal of the American Medical Informatics Association : JAMIA*, 19(3):328–33, 2012.

[31] D. N. Wright. Does a post-take ward round proforma have a positive effect on completeness of documentation and efficiency of information management? *Health informatics journal*, 15(2):86–94, 2009.

[32] D. Kalra. Electronic health record standards. *Methods of Information in Medicine*, 45:136–144, 2006.

[33] Kuhn KA Lenz R. Integration of heterogeneous and autonomous systems in hospitals. *Business Brie ng: Data management & Storage Technology*, 2002.

[34] A. Edwards, I. Hollin, J. Barry, and S. Kachnowski. Barriers to cross–institutional health information exchange: a literature review. *Journal of healthcare information management : JHIM*, 24(3):22–34, 2010.

[35] A. Hoerbst and E. Ammenwerth. Electronic health records a systematic review on quality requirements. *Methods of Information in Medicine*, 49(4):320–336, 2010.

[36] R. Haux. Health information systems - past, present, future. *International journal of medical informatics*, 75(3-4):268–81, 2006.

[37] European Comission. European hospital survey - benchmarking deployment of ehealth services (2012-2013). Technical report, EU, 2014.

[38] EU. Communication from the commission to the council, the european parliament, the european economic and social committee and the committee of the regions : e-health - making healthcare better for european citizens: An action plan for a european e-health area. Technical report, EU, 2004.

[39] EU. Commission recommendation of 2 july 2008 on cross-border interoperability of electronic health record systems. Technical report, EU, 2008.

[40] epsos project. http://www.epsos.eu.

[41] European Comission. ehealth action plan 2012-2020 - innovative healthcare for the 21st century. Technical report, European Comission, 2012.

[42] empirica GmbH. Ict standards in the health sector: current situation and prospects,. http://ec.europa.eu/enterprise/archives/e-business-watch/studies/special_topics/2007/documents/Special-study_01-2008_ICT_health_standards.pdf, 2008.

[43] Beale Thomas. Archetypes: constraint-based domain models for future-proof information systems. In *Proceedings of the 11th OOPSLA Workshop on Behavioral Semantics: Serving the Customer*, page pp. 16ñ32., 2002.

[44] J. Bisbal and D. Berry. An analysis framework for electronic health record systems. interoperation and collaboration in shared healthcare. *Methods Inf Med*, 50(2):180–9, 2011.

[45] B. G. Blobel, K. Engel, and P. Pharow. Semantic interoperability–hl7 version 3 compared to advanced architecture standards. *Methods Inf Med*, 45(4):343–53, 2006.

[46] J. Grimson, W. Grimson, D. Berry, G. Stephens, E. Felton, D. Kalra, P. Toussaint, and O. W. Weier. A corba-based integration of distributed electronic healthcare records using the synapses approach. *IEEE Trans Inf Technol Biomed*, 2(3):124–38, 1998.

[47] Peter Schloeffel, Thomas Beale, George Hayworth, Sam Heard, and Heather Leslie. The relationship between cen 13606, hl7, and openehr. In *Procedings of HIC 2006 bridging the digital divide: clinician, consumer and computer*, 2006.

[48] Snomed ct. http://ihtsdo.org/fileadmin/user_upload/doc/.

[49] ICD - International Classification of Diseases. http://www.cdc.gov/nchs/icd.htm.

[50] Michael Wooldridge and Nicholas R. Jennings. Intelligent agents: Theory and practice. *Knowledge Engineering Review*, 10(2):115–152, 1995.

[51] FIPA. Fipa Agent Management Specification. http://www.fipa.org /specs/fipa00023/XC00023H.html.

[52] Hong Lin. *Architectural Design of Multi-Agent Systems: Technologies and Techniques*. IGI-Global, 2007.

[53] John Nealon and Antonio Moreno. Agent-based applications in health care. In *Applications of Software Agent Technology in the Health Care Domain , Whitestein Series in Software Agent Technologies, Birkhauser Verlag*, pages 3–18. Verlag, 2003.

[54] D. Isern, D. Sanchez, and A. Moreno. Agents applied in health care: A review. *International journal of medical informatics*, 79(3):145–66, 2010.

[55] IlKon Kim and JiHyun Yun. Agent-based intelligent clinical information system for persistent lifelong electronic medical record. In Jaeho Lee and Mike Barley, editors, *Intelligent Agents and Multi-Agent Systems*, volume 2891 of *Lecture Notes in Computer Science*, pages 194–204. Springer Berlin Heidelberg, 2003.

[56] António Abelha, José Machado, Victor Alves, and José Neves. Health data management in the medical arena. In *Proceedings of the 4th WSEAS International Conference on Applied Informatics and Communications*, AIC'04, pages 14:1–14:6, Stevens Point, Wisconsin, USA, 2004.

[57] B. Orgun and J. Vu. Hl7 ontology and mobile agents for interoperability in heterogeneous medical information systems. *Computers in Biology and Medicine*, 36(7-8):817–836, 2006.

[58] J. Sayyad Shirabad, S. Wilk, W. Michalowski, and K. Farion. Implementing an integrative multi-agent clinical decision support system with open source software. *J Med Syst*, 36(1):123–37, 2012.

[59] Barbara A. Kitchenham, Tore Dyba, and Magne Jorgensen. Evidence-based software engineering. In *Proceedings of the 26th International Conference on Software Engineering*, ICSE '04, pages 273–281, Washington, DC, USA, 2004.

[60] Green S . Higgins JPT, editor. *Cochrane Handbook for Systematic Reviews of Interventions*. The Cochrane Collaboration, 2011.

[61] Jorge Biolchini, Paula Gomes Mian, Ana Candida Cruz Natali, and Guilherme Horta Travassos. Systematic review in software engineering. Technical report, Systems Engineering and Computer Science Department, COPPE/UFRJ, 2005.

[62] David Moher, Alessandro Liberati, Jennifer Tetzlaff, Douglas G. Altman, and PRISMA Grp. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *INTERNATIONAL JOURNAL OF SURGERY*, 8(5):336–341, 2010.

[63] Jorge Jacome. Medquest - um projecto de qquestion·rios web n ·rea da sa de. Master's thesis, Faculty of Medicine, University of Porto, 2009.

[64] Michael Wooldridge, Nicholas R. Jennings, and David Kinny. The gaia methodology for agent-oriented analysis and design. *Autonomous Agents and Multi-Agent Systems*, 3(3):285–312, September 2000.

[65] Lin Padgham and Michael Winikoff. Prometheus: A methodology for developing intelligent agents. In *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems: Part 1*, AAMAS '02, pages 37–38, New York, NY, USA, 2002.

[66] ScottA. DeLoach. The mase methodology. In Federico Bergenti, Marie-Pierre Gleizes, and Franco Zambonelli, editors, *Methodologies and Software Engineering for Agent Systems*, volume 11 of *Multiagent Systems, Arti cial Societies, and Simulated Organizations*, pages 107–125. Springer US, 2004.

[67] Carles Sierra. Electronics institutions: Methodology of multi-agent systems development. In Chengqi Zhang, Hans W. Guesgen, and Wai-Kiang Yeap, editors, *PRICAI 2004: Trends in Arti cial Intelligence*, volume 3157 of *Lecture Notes in Computer Science*, pages 2–2. Springer Berlin Heidelberg, 2004.

[68] Marc-Philippe Huget, James Odell, and Bernhard Bauer. The auml approach. In Federico Bergenti, Marie-Pierre Gleizes, and Franco Zambonelli, editors, *Methodologies and Software Engineering for Agent Systems*, volume 11 of *Multiagent Systems, Arti cial Societies, and Simulated Organizations*, pages 237–257. Springer US, 2004.

[69] Umls - Unified Medical Language Systems. www.nlm.nih.gov/research/umls/.

[70] Fabio Bellifemine, Agostino Poggi, and Giovanni Rimassa. Jade: A fipa2000 compliant agent development environment. In *Proceedings of the Fifth International Conference on Autonomous Agents*, AGENTS '01, pages 216–217, New York, NY, USA, 2001.

[71] Nick Howden, Ralph Rönnquist, Andrew Hodgson, and Andrew Lucas. Jack intelligent agents-summary of an agent infrastructure. In *5th International conference on autonomous agents*, 2001.

[72] Danny B Lange, Mitsuru Oshima, Günter Karjoth, and Kazuya Kosaka. Aglets: Programming mobile agents in java. In *Worldwide Computing and Its Applications*, pages 253–266. Springer, 1997.

[73] Marcela Rodríguez and Jesus Favela. A framework for supporting autonomous agents in ubiquitous computing environments. In *Proceedings of System Support for Ubiquitous Computing Workshop at the Fifth Annual Conference on Ubiquitous Computing (UbiComp 2003), Seattle, Washington*, 2003.

[74] Massimo Zaniboni. Jude framework. http://jude.sourceforge.net/, 2001.

[75] Olivier Gutknecht and Jacques Ferber. The madkit agent platform architecture. In *Revised Papers from the International Workshop on Infrastructure for Multi-Agent Systems: Infrastructure for Agents, Multi-Agent Systems, and Scalable Multi-Agent Systems*, pages 48–55, London, UK, UK, 2001.

[76] John E. Laird, Allen Newell, and Paul S. Rosenbloom. Soar: An architecture for general intelligence. *Arti cial Intelligence*, 33(1):1 – 64, 1987.

[77] Crepeau R. Fipa smart v3.0 - fipa-based stationary and mobile agent resource toolkit.

[78] William J Clancey, Patricia Sachs, Maarten Sierhuis, and RON VAN HOOF. Brahms: Simulating practice for work systems design. *International Journal of Human-Computer Studies*, 49(6):831–865, 1998.

[79] David L Martin, Adam J Cheyer, and Douglas B Moran. The open agent architecture: A framework for building distributed software systems. *Applied Arti cial Intelligence*, 13(1-2):91–128, 1999.

[80] Repast simphony 2.2. http://repast.sourceforge.net/.

[81] C. S. Wu, I. Khoury, and H. Shah. Optimizing medical data quality based on multiagent web service framework. *Ieee Transactions on Information Technology in Biomedicine*, 16(4):745–757, 2012.

[82] A. Ab Aziz, M. C. A. Klein, and J. Treur. An integrative ambient agent model for unipolar depression relapse prevention. *Journal of Ambient Intelligence and Smart Environments*, 2(1):5–20, 2010.

[83] D. R. Fernandez, J. M. G. Chamizo, F. M. Parez, and A. S. Paya. *Modelling of dysfunctions in the neuronal control of the lower urinary fact*, volume 3561 of *Lecture Notes in Computer Science*, pages 203–212. Springer-Verlag Berlin, Berlin, 2005.

[84] E. Tello-Leal, O. Chiotti, and P. D. Villarreal. Process-oriented integration and coordination of healthcare services across organizational boundaries. *Journal of Medical Systems*, 36(6):3713–3724, 2012.

[85] M. V. Sokolova and A. Fernandez-Caballero. Modeling and implementing an agent-based environmental health impact decision support system. *Expert Systems with Applications*, 36(2):2603–2614, 2009.

[86] M. V. Sokolova and A. Fernandez-Caballero. Hybrid models in agent-based environmental decision support. *Applied Soft Computing*, 11(8):5243–5258, 2011.

[87] M. Druzovec, T. Welzer, and B. Brumen. Agent oriented approach to handling medical data. *J Med Syst*, 29(1):45–57, 2005.

[88] C. J. Su and C. W. Peng. Multi-agent ontology-based web 2.0 platform for medical rehabilitation. *Expert Systems with Applications*, 39(12):10311–10323, 2012.

[89] H. Gonzalez-Velez, M. Mier, M. Julia-Sape, T. N. Arvanitis, J. M. Garcia-Gomez, M. Robles, P. H. Lewis, S. Dasmahapatra, D. Dupplaw, A. Peet, C. Arus, B. Celda, S. Van Huffel, and M. Lluch-Ariet. Healthagents: distributed multi-agent brain tumor diagnosis and prognosis. *Applied Intelligence*, 30(3):191–202, 2009.

[90] P. Bloodsworth and S. Greenwood. Cosmoa an ontology-centric multi-agent system for co-ordinating medical responses to large-scale disasters. *Ai Communications*, 18(3):229–240, 2005.

[91] F. Bouchoul and M. Mostefai. Agent-services and mobile agents for an integrated hcis. *International Journal of Computer Integrated Manufacturing*, 22(5):458–471, 2009.

[92] G. E. Phillips-Wren and G. A. Forgionne. *Intelligent decision making in information retrieval*, volume 3213 of *Lecture Notes in Computer Science*, pages 103–109. Springer-Verlag Berlin, Berlin, 2004.

[93] N. Paoletti, P. Lio, E. Merelli, and M. Viceconti. Multilevel computational modeling and quantitative analysis of bone remodeling. *Computational Biology and Bioinformatics, IEEE/ACM Transactions on*, 9(5):1366–1378, 2012.

[94] F. Pappalardo, S. Musumeci, and S. Motta. Modeling immune system control of atherogenesis. *Bioinformatics*, 24(15):1715–1721, 2008.

[95] D. Gritzalis and C. Lambrinoudakis. A security architecture for interconnecting health information systems. *International Journal of Medical Informatics*, 73(3):305–309, 2004.

[96] K. Nouira and A. Trabelsi. Intelligent monitoring system for intensive care units. *Journal of Medical Systems*, 36(4):2309–2318, 2012.

[97] B. Lopez, B. Innocenti, and D. Busquets. A multiagent system for coordinating ambulances for emergency medical services. *Intelligent Systems, IEEE*, 23(5):50–57, 2008.

[98] E. G. P. Bovenkamp, J. Dijkstra, J. G. Bosch, and J. H. C. Reiber. Multi-agent segmentation of ivus images. *Pattern Recognition*, 37(4):647–663, 2004.

[99] M. Nikooghadam and A. Zakerolhosseini. Secure communication of medical information using mobile agents. *Journal of Medical Systems*, 36(6):3839–3850, 2012.

[100] J. Choe and S. K. Yoo. Web-based secure access from multiple patient repositories. *International Journal of Medical Informatics*, 77(4):242–248, 2008.

[101] S. Walczak. A multiagent architecture for developing medical information retrieval agents. *J Med Syst*, 27(5):479–98, 2003.

[102] M. McShane, S. Beale, S. Nirenburg, B. Jarrell, and G. Fantry. Inconsistency as a diagnostic tool in a society of intelligent agents. *Arti cial Intelligence in Medicine*, 55(3):137–148, 2012.

[103] V. Chan, P. Ray, and N. Parameswaran. Mobile e-health monitoring: an agent-based approach. *Communications, IET*, 2(2):223–230, 2008.

[104] V. Rialle, J. B. Lamy, N. Noury, and L. Bajolle. Telemonitoring of patients at home: a software agent approach. *Computer Methods and Programs in Biomedicine*, 72(3):257–268, 2003.

[105] D. Isern, D. Sanchez, and A. Moreno. Ontology-driven execution of clinical guidelines. *Computer Methods and Programs in Biomedicine*, 107(2):122–139, 2012.

[106] D. Isern, A. Valls, and A. Moreno. *Using aggregation operators to personalize agent-based medical services*, volume 4252 of *Lecture Notes in Arti cial Intelligence*, pages 1256–1263. Springer-Verlag Berlin, Berlin, 2006.

[107] A. Moreno, A. Valls, D. Isern, and D. Sanchez. Applying agent technology to healthcare: The grusma experience. *Intelligent Systems, IEEE*, 21(6):63–67, 2006.

[108] J. C. D. Vazquez, A. C. Martinez, A. Gomez, and B. A. Varela. Intelligent agents technology applied to tasks scheduling and communications management in a critical care telemonitoring system. *Computers in Biology and Medicine*, 37(6):760–773, 2007.

[109] A. Moreno, D. Isern, and D. Sanchez. Provision of agent-based health care services. *Ai Communications*, 16(3):167–178, 2003.

[110] J. M. Corchado, D. I. Tapia, and J. Bajo. A multi-agent architecture for distributed services and applications. *International Journal of Innovative Computing Information and Control*, 8(4):2453–2476, 2012.

[111] J. M. Corchado, J. Bajo, and A. Abraham. Gerami: Improving healthcare delivery in geriatric residences. *Intelligent Systems, IEEE*, 23(2):19–25, 2008.

[112] M. Ndiaye, J. J. Montois, and A. Kinie. Intelligent control in vectorial analysis of epileptic signals: a multi-agents experimental approach. *Traitement Du Signal*, 24(6):447–459, 2007.

[113] J. J. Chen, P. W. Wang, Y. C. Huang, and H. C. Yen. Applying ontology techniques to develop a medication history search and alert system in department of nuclear medicine. *Journal of Medical Systems*, 36(2):737–746, 2012.

[114] T. L. Chen, Y. F. Chung, and F. Y. S. Lin. Deployment of secure mobile agents for medical information systems. *Journal of Medical Systems*, 36(4):2493–2503, 2012.

[115] C. H. Liu, Y. F. Chung, T. S. Chen, and S. D. Wang. Mobile agent application and integration in electronic anamnesis system. *Journal of Medical Systems*, 36(3):1009–1020, 2012.

[116] C. H. Liu, Y. F. Chung, T. W. Chiang, T. S. Chen, and S. D. Wang. A mobile agent approach for secure integrated medical information systems. *Journal of Medical Systems*, 36(5):2731–2741, 2012.

[117] C. S. Lee and M. H. Wang. Ontology-based intelligent healthcare agent and its application to respiratory waveform recognition. *Expert Systems with Applications*, 33(3):606–619, 2007.

[118] L. Godo, J. Puyol-Gruart, J. Sabater, V. Torra, P. Barrufet, and X. Fabregas. A multi-agent system approach for monitoring the prescription of restricted use antibiotics. *Arti cial Intelligence in Medicine*, 27(3):259–282, 2003.

[119] L. Catarinucci, R. Colella, A. Esposito, L. Tarricone, and M. Zappatore. Rfid sensor-tags feeding a context-aware rule-based healthcare monitoring system. *Journal of Medical Systems*, 36(6):3435–3449, 2012.

[120] G. Koum, A. Yekel, B. Ndifon, J. Etang, and F. Simard. Design of a two-level adaptive multi-agent system for malaria vectors driven by an ontology. *BMC Med Inform Decis Mak*, 7:19, 2007.

[121] T. Alsinet, C. Ansotegui, R. Bejar, C. Fernandez, and F. Manya. Automated monitoring of medical protocols: a secure and distributed architecture. *Arti cial Intelligence in Medicine*, 27(3):367–392, 2003.

[122] X. B. Yang and Y. Miao. Distributed agent based interoperable virtual emr system for healthcare system integration. *Journal of Medical Systems*, 35(3):309–319, 2011.

[123] Y. Ji, H. Ying, J. Yen, S. Zhu, D. C. Barth-Jones, R. E. Miller, and R. M. Massanari. A distributed adverse drug reaction detection system using intelligent agents with a fuzzy recognition-primed decision model. *International Journal of Intelligent Systems*, 22(8):827–845, 2007.

[124] V. Della Mea. Agents acting and moving in healthcare scenario - a paradigm for telemedical collaboration. *Ieee Transactions on Information Technology in Biomedicine*, 5(1):10–13, 2001.

[125] M. Tentori, M. D. Rodriguez, and J. Favela. An agent-based middleware for the design of activity-aware applications. *Ieee Intelligent Systems*, 26(3):15–23, 2011.

[126] M. Tentori, J. Favela, and M. D. Rodriguez. Privacy-aware autonomous agents for pervasive healthcare. *Intelligent Systems, IEEE*, 21(6):55–62, 2006.

[127] M. D. Rodriguez, J. Favela, A. Preciado, and A. Vizcaino. Agent-based ambient intelligence for healthcare. *Ai Communications*, 18(3):201–216, 2005.

[128] J. Favela, M. Rodriguez, A. Preciado, and V. M. Gonzalez. Integrating context-aware public displays into a mobile hospital information system. *Information Technology in Biomedicine, IEEE Transactions on*, 8(3):279–286, 2004.

[129] M. Rodriguez and A. Preciado. *An agent based system for the contextual retrieval of medical information*, volume 3034 of *Lecture Notes in Arti cial Intelligence*, pages 64–73. Springer-Verlag Berlin, Berlin, 2004.

[130] I. C. Oliveira, O. Belo, and J. P. Cunha. *Agents working on the integration of heterogeneous information sources in distributed healthcare environments*, volume 1952 of *Lecture Notes in Arti cial Intelligence*, pages 136–145. Springer-Verlag Berlin, Berlin, 2000.

[131] C. J. Su and C. Y. Wu. Jade implemented mobile multi-agent based, distributed information platform for pervasive health care monitoring. *Applied Soft Computing*, 11(1):315–325, 2011.

[132] C. J. Su. Mobile multi-agent based, distributed information platform (madip) for wide-area e-health monitoring. *Computers in Industry*, 59(1):55–68, 2008.

[133] M. Ulieru, M. Hadzic, and E. Chang. Soft computing agents for e-health in application to the research and control of unknown diseases. *Information Sciences*, 176(9):1190–1214, 2006.

[134] C. C. Marinagi, C. D. Spyropoulos, C. Papatheodorou, and S. Kokkotos. Continual planning and scheduling for managing patient tests in hospital laboratories. *Arti cial Intelligence in Medicine*, 20(2):139–154, 2000.

[135] L. D. Serbanati, F. L. Ricci, G. Mercurio, and A. Vasilateanu. Steps towards a digital health ecosystem. *Journal of Biomedical Informatics*, 44(4):621–636, 2011.

[136] P. Tolchinsky, U. Cortes, S. Modgil, F. Caballero, and A. Lopez-Navidad. Increasing human-organ transplant availability: Argumentation-based agent deliberation. *Intelligent Systems, IEEE*, 21(6):30–37, 2006.

[137] J. Vazquez-Salceda, J. A. Padget, U. Cortes, A. Lopez-Navidad, and F. Caballero. Formalizing an electronic institution for the distribution of human tissues. *Arti cial Intelligence in Medicine*, 27(3):233–258, 2003.

[138] J. Vazquez-Salceda, U. Cortes, J. Padget, A. Lopez-Navidad, and F. Caballero. The organ allocation process: a natural extension of the carrel agent-mediated electronic institution. *Ai Communications*, 16(3):153–165, 2003.

[139] K. Decker and J. J. Li. Coordinating mutually exclusive resources using gpgp. *Autonomous Agents and Multi-Agent Systems*, 3(2):133–157, 2000.

[140] J. T. Murphy, R. Walshe, and M. Devocelle. A theoretical analysis of the prodrug delivery system for treating antibiotic-resistant bacteria. *Computational Biology and Bioinformatics, IEEE/ACM Transactions on*, 8(3):650–658, 2011.

[141] G. Lanzola, L. Gatti, S. Falasconi, and M. Stefanelli. A framework for building cooperative software agents in medical applications. *Arti cial Intelligence in Medicine*, 16(3):223–249, 1999.

[142] G. Lanzola, M. Campagnoli, S. Falasconi, and M. Stefanelli. *A framework for building cooperating agents*, volume 1211 of *Lecture Notes in Arti cial Intelligence*, pages 479–490. Springer-Verlag Berlin, Berlin, 1997.

[143] G. Lanzola, S. Falasconi, and M. Stefanelli. *Cooperative software agents for patient management*, volume 934 of *Arti cial Intelligence in Medicine*. Springer-Verlag Berlin, Berlin 33, 1995.

[144] M. Laskowski, B. C. P. Demianyk, J. Witt, S. N. Mukhi, M. R. Friesen, and R. D. McLeod. Agent-based modeling of the spread of influenza-like illness in an emergency department: A simulation study. *Information Technology in Biomedicine, IEEE Transactions on*, 15(6):877–889, 2011.

[145] T. Kifor, L. Z. Varga, J. Vazquez-Salceda, S. Alvarez, S. Willmott, S. Miles, and L. Moreau. Provenance in agent-mediated healthcare systems. *Intelligent Systems, IEEE*, 21(6):38–46, 2006.

[146] L. M. Camarinha-Matos and W. Vieira. Intelligent mobile agents in elderly care. *Robotics and Autonomous Systems*, 27(1-2):59–75, 1999.

[147] D. Isern, A. Moreno, D. Sanchez, A. Hajnal, G. Pedone, and L. Z. Varga. Agent-based execution of personalised home care treatments. *Applied Intelligence*, 34(2):155–180, 2011.

[148] R. Hill, S. Polovina, and D. Shadija. *Transaction agent modelling: From experts to concepts to multi-agent systems*, volume 4068 of *Lecture Notes in Arti cial Intelligence*, pages 247–259. Springer-Verlag Berlin, Berlin, 2006.

[149] B. G. Silverman. The role of web agents in medical knowledge management. *M D Computing*, 15(4):221–+, 1998.

[150] A. G. Dunn, M. S. Ong, J. I. Westbrook, F. Magrabi, E. Coiera, and W. Wobcke. A simulation framework for mapping risks in clinical processes: the case of in-patient transfers. *Journal of the American Medical Informatics Association*, 18(3):259–266, 2011.

[151] J. Fox, D. Glasspool, and S. Modgil. A canonical agent model for healthcare applications. *Intelligent Systems, IEEE*, 21(6):21–28, 2006.

[152] S. N. Murphy, T. Ng, D. F. Sittig, and G. O. Barnett. Using web technology and java mobile software agents to manage outside referrals. *Journal of the American Medical Informatics Association*, pages 101–105, 1998.

[153] P. De Meo, G. Quattrone, and D. Ursino. Integration of the hl7 standard in a multiagent system to support personalized access to e-health services. *Knowledge and Data Engineering, IEEE Transactions on*, 23(8):1244–1260, 2011.

[154] C. Caceres, A. Fernandez, S. Ossowski, and M. Vasirani. Agent-based semantic service discovery for healthcare: An organizational approach. *Intelligent Systems, IEEE*, 21(6):11–20, 2006.

[155] J. E. Larssan and B. Hayes-Roth. Guardian: intelligent autonomous agent for medical monitoring and diagnosis. *Intelligent Systems and their Applications, IEEE*, 13(1):58–64, 1998.

[156] B. HayesRoth and J. E. Larsson. A domain-specific software architecture for a class of intelligent patient monitoring agents. *Journal of Experimental & Theoretical Arti cial Intelligence*, 8(2):149–171, 1996.

[157] B. Hayesroth. An architecture for adaptive intelligent systems. *Arti cial Intelligence*, 72(1-2):329–365, 1995.

[158] B. Hayesroth, S. Uckun, J. E. Larsson, J. Drakopoulos, D. Gaba, J. Barr, and J. Chien. Guardian - an experimental system for intelligent icu monitoring. *Journal of the American Medical Informatics Association*, pages 1004–1004, 1994.

[159] Ji Yanqing, Ying Hao, M. S. Farber, J. Yen, P. Dews, R. E. Miller, and R. M. Massanari. A distributed, collaborative intelligent agent system approach for proactive postmarketing drug safety surveillance. *Information Technology in Biomedicine, IEEE Transactions on*, 14(3):826–837, 2010.

[160] J. Vazquez-Salceda. Normative agents in health care: Uses and challenges. *Ai Communications*, 18(3):175–189, 2005.

[161] S. Falasconi, L. Dazzi, G. Lanzola, S. Quaglini, R. Saracco, and M. Stefanelli. Towards cooperative patient management through organizational and medical ontologies. *Methods of Information in Medicine*, 37(4-5):564–575, 1998.

[162] S. Falasconi, G. Lanzola, and M. Stefanelli. Ontology and terminology servers in agent-based health-care information systems. *Methods of Information in Medicine*, 36(1):30–43, 1997.

[163] S. Falasconi, G. Lanzola, and M. Stefanelli. An ontology-based multi-agent architecture for distributed health-care information systems. *Methods of Information in Medicine*, 36(1):20–29, 1997.

[164] Xu Xuyan and Li Lefei. An artificial urban healthcare system and applications. *Intelligent Systems, IEEE*, 25(3):63–73, 2010.

[165] C. Reed, B. Boswell, and R. Neville. *Multi-agent patient representation in primary care*, volume 3581 of *Lecture Notes in Arti cial Intelligence*, pages 375–384. Springer-Verlag Berlin, Berlin, 2005.

[166] S. N. Murphy, U. H. Rabbani, and G. O. Barnett. Using software agents to maintain autonomous patient registries for clinical research. *Journal of the American Medical Informatics Association*, pages 71–75, 1997.

[167] L. Xiao, B. Hu, M. Croitoru, P. Lewis, and S. Dasmahapatra. A knowledgeable security model for distributed health information systems. *Computers & Security*, 29(3):331–349, 2010.

[168] I. A. Letia and M. Acalovschi. *Achieving competence by argumentation on rules for roles*, volume 3451 of *Lecture Notes in Arti cial Intelligence*, pages 45–59. Springer-Verlag Berlin, Berlin, 2005.

[169] A. Hutchison, M. Kaiserswerth, M. Moser, and A. Schade. Electronic data interchange for health care. *Communications Magazine, IEEE*, 34(7):28–34, 1996.

[170] D. Vassis, P. Belsis, C. Skourlas, and G. Pantziou. Providing advanced remote medical treatment services through pervasive environments. *Personal and Ubiquitous Computing*, 14(6):563–573, 2010.

[171] V. G. Koutkias, I. Chouvarda, and N. Maglaveras. A multiagent system enhancing home-care health services for chronic disease management. *Information Technology in Biomedicine, IEEE Transactions on*, 9(4):528–537, 2005.

[172] J. Huang, N. R. Jennings, and J. Fox. Agent-based approach to health-care management. *Applied Arti cial Intelligence*, 9(4):401–420, 1995.

[173] J. A. Fraile, J. Bajo, J. M. Corchado, and A. Abraham. Applying wearable solutions in dependent environments. *Information Technology in Biomedicine, IEEE Transactions on*, 14(6):1459–1467, 2010.

[174] A. Hara, T. Ichimura, T. Takahama, and Y. Isomichi. *Extraction of risk factors by multi-agent voting model using automatically de ned groups*, volume 3683 of *Lecture Notes in Arti cial Intelligence*, pages 1218–1224. Springer-Verlag Berlin, Berlin, 2005.

[175] A. Hara, T. Ichimura, T. Takahama, and Y. Isomichi. *Extraction of rules from coronary heart disease database using automatically de ned groups*, volume 3214 of *Lecture Notes in Computer Science*, pages 1089–1096. Springer-Verlag Berlin, Berlin, 2004.

[176] N. Bonjean, C. Bernon, and P. Glize. Towards a guide for engineering the collective behaviour of a mas. *Simulation Modelling Practice and Theory*, 18(10):1506–1514, 2010.

[177] T. Guyet, C. Garbay, and M. Dojat. *Human/computer interaction to learn scenarios from ICU multivariate time series*, volume 3581 of *Lecture Notes in Arti cial Intelligence*, pages 424–428. Springer-Verlag Berlin, Berlin, 2005.

[178] Nagios. http://www.nagios.org/.

[179] Network working group of Cisco Systems. Us secure hash algorithm 1 (sha1). rfc 3174. ftp://ftp-rfc-editor.org/in-notes/rfc2246.txt, 1999.

[180] Nomenclatura das unidade territoriais para fins estatÌsticos. http://smi.ine.pt/Conceito/Detalhes/3879?modal=1, 2004.

[181] PORDATA. http://www.pordata.pt/Municipios/Populacao+ residente+ por+migracoes+segundo+os+Censos-48.

[182] D Akkerman, L Cleland, G Croft, K Eskuchen, C Heim, A Levine, L Setterlund, C Stark, J Vickers, and E Westby. Routine prenatal care. Technical report, Institute for Clinical Systems Improvement., 2012.

[183] D Creedon, L Atwood, L Bates, C Rose, A Levin, C McCall, R Wingeier, B Walkes, D Akkerman, and L Pratt. Health care guideline: Management of labor, 4th ed. Technical report, Institute for Clinical Systems Improvement., 2011.

[184] Jason S. Shapiro, Joseph Kannry, Andre W. Kushniruk, Gilad Kuperman, and NYCLIX Clincal Advisory Subcomm. Emergency physicians' perceptions of health information exchange. *J. Am. Med. Inform. Assoc.*, 14(6):700–705, 2007.

[185] J. Ametller, S. Robles, and J. A. Ortega-Ruiz. Self-protected mobile agents. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems - Volume 1*, AAMAS '04, pages 362–367, Washington, DC, USA, 2004. IEEE Computer Society.

[186] G. Karjoth, N. Asokan, and C. Guelcue. Protecting the computation results of free-roaming agents. *Personal Technologies*, 2(2):92–99, 1998.

[187] Apache tomcat. http://tomcat.apache.org/.

[188] existdb. http://exist-db.org/exist/apps/homepage/index.html.

[189] Erik Sundvall, Mikael Nystrom, Daniel Karlsson, Martin Eneling, Rong Chen, and Hakan Orman. Applying representational state transfer (rest) architecture to archetype-based electronic health record systems. *BMC Medical Informatics and Decision Making*, 13(1), 2013.

[190] Abraham MartÌn-Campillo, Carles MartÌnez-GarcÌa, Jordi Cucurull, Ramon MartÌ, Sergi Robles, and Joan Borrell. Mobile agents in healthcare, a distributed intelligence approach. In Isabelle Bichindaritz, Sachin Vaidya, Ashlesha Jain, and LakhmiC. Jain, editors, *Computational Intelligence in Healthcare 4*, volume 309 of *Studies in Computational Intelligence*, pages 49–80. Springer Berlin Heidelberg, 2010.

A QUESTIONNAIRE USED FOR PROFILING DATA NEEDS

This questionnaire was used for gathering data among health professional with the purpose of characterising different profiles of health data usage. It includes three sections : the health professional identification (demographics and type os specialisation); the characterisation of the current status with regard to access to information from other institutions and the characterisation of information needs in different settings of care (Emergency, Inpatient, Outpatient).

The questionnaire was applied in portuguese, for clarity and usefulness the version that is made available in the next pages was translated into english.

Nº:
Date: ___/___/____
Place:

[illegible]

Given an increasingly common scenario geographical mobility of users, we plan to evaluate with this questionnaire among health professionals, in various contexts of health care, the needs of clinical information from other health.

I - [illegible]

Age: ☐

Gender: Masc. ☐ Femin. ☐

Medical Speciality: ____

Do you have access to a computer in your work settings? Yes ☐ No ☐

From the last 100 patients that you observed what was the proportion in the following contexto of care.:

Emergency	Inpatient	Outpatient
___ /100	___ /100	___ /100

For data types in the following table, indicate those that, in your institution, are accessible via a Computer Information System:

General and Administrative Data	Yes	No
Data for Alerts	Yes	No
Data from Diagnosis	Yes	No
Data from Pharmacological Treatments	Yes	No
Data from Surgical Treatments	Yes	No
Data from Inpatient Events	Yes	No
Data from Outpatient Events	Yes	No
Data from Day Hospital Events	Yes	No
Data from Emergency Events	Yes	No
Data from Laboratory Reports	Yes	No
Data from Iamgiology.	Yes	No
Data from Discharge Reports	Yes	No
Nursing Notes	Yes	No
Data from Surgery	Yes	No
Patient Managed Data (Pregnancy Bulletin/ Vaccination etc.)	Yes	No

II - Characterization of the current status with regard to access to information from other institutions

P1 – How important is the ability to access clinical data from other institutions?

Unimportant	1	2	3	4	5	Very Important

P2 – How do you rate the access to clinical data from other institutions in your workplace?

Easy	1	2	3	4	5	Dificult

P3 – What percentage of cases in which attempts to obtain information about their patients outside your institution
?

0-25	25-50	50-75	75-100

P4 – While trying to access external information of your patients, which are the primary means of communication that you use? (Select only one)

Information System	
e-Mail	
Telephone	
Mail	

P5 – Order in accordance with the importance given to each source of (1 – Unimportant a 4 – Very Important)?

Private Clinics	
Primary Care	
Private Labs	
Hospitals	

P6 – How important is the ability to access clinical data in digital media?

Unimportant	1	2	3	4	5	Very Important

P7 – Indicate a factor that, in your opinion, is an obstacle to the exchange of clinical information between health institutions: ____

P8 – The availability of information from other institutions would benefit: (Put a cross in each line)

	Totally Agree	Agree	Indifferent	Disagree	Totally Disagree
Your performance as an health professional					
Your department					
Your institution					
The National Health Service as a hole					

P9 – The availability of information from other institutions would increase or decrease … : (Put a cross in each line)

	Would largely Increase	Would Increase	No impact	Would be reduced	Would be largely reduced
The efficiency in the provision of care					
The number of tests requested					
The number of medical errors					
the time available to make decisions about the patient.					

III – Information Needs Characterization

P1 - According to the scale outlined, taking into account your experience / expertise indicate the importance of having access to various types of data from other health institutions in the various contexts of care provision:

" 0 " - Never Important;
" 1 " – Sometimes Important;
" 2 " – Always Important:

Example			
Type of Data	Emergency	Inpatient	Outpatient
Never important in *Emergency,* ***Sometimes Important*** *in Inpatient,* ***Always important*** *in Outpatient*	X 1 2	0 X 2	0 1 X

Type of Data	Emergency	Inpatient	Outpatient
General Data			
Patient Identification	0 1 2	0 1 2	0 1 2
Gender	0 1 2	0 1 2	0 1 2
Name	0 1 2	0 1 2	0 1 2
Age	0 1 2	0 1 2	0 1 2
Address	0 1 2	0 1 2	0 1 2
Marital Status	0 1 2	0 1 2	0 1 2
Profession	0 1 2	0 1 2	0 1 2
Phone Number	0 1 2	0 1 2	0 1 2
Blood type	0 1 2	0 1 2	0 1 2
Alerts			
Allergies (Adverse Reactions and Sensibility)	0 1 2	0 1 2	0 1 2
Diabetes	0 1 2	0 1 2	0 1 2

Type of Data	Emergency	Inpatient	Outpatient
Diagnosis			
Infectious and Parasitic Diseases	0 1 2	0 1 2	0 1 2
Neoplasms	0 1 2	0 1 2	0 1 2
Endocrine, nutritional and metabolic diseases and immunity disorders	0 1 2	0 1 2	0 1 2
Diseases of the Blood and Hematopoietic Organs	0 1 2	0 1 2	0 1 2
Mental Disorders	0 1 2	0 1 2	0 1 2
Nervous System and Sense Organs diseases	0 1 2	0 1 2	0 1 2
Circulatory System Diseases	0 1 2	0 1 2	0 1 2
Respiratory Diseases	0 1 2	0 1 2	0 1 2
Digestive System Diseases	0 1 2	0 1 2	0 1 2
Genitourinary System Diseases	0 1 2	0 1 2	0 1 2
Complications of Pregnancy, Childbirth and the Puerperium	0 1 2	0 1 2	0 1 2
Skin and Subcutaneous Tissue Cell Diseases	0 1 2	0 1 2	0 1 2
Musculoskeletal System and Connective Tissue Diseases	0 1 2	0 1 2	0 1 2
Congenital Anomalies	0 1 2	0 1 2	0 1 2
Some disorders Originating in Perinatal Period	0 1 2	0 1 2	0 1 2
Symptoms, signs and ill-defined conditions	0 1 2	0 1 2	0 1 2
Injuries and Poisoning	0 1 2	0 1 2	0 1 2
Treatments			
Pharmacological (Current Medication, Past Prescription, Desensitization)	0 1 2	0 1 2	0 1 2
Surgical	0 1 2	0 1 2	0 1 2
Rehabilitation (ex: physiatric)	0 1 2	0 1 2	0 1 2
Homeopathic	0 1 2	0 1 2	0 1 2
Events			
Inpatient Data (Date, Reason, Department, Outcome, Destination after discharge)	0 1 2	0 1 2	0 1 2
Outpatients Data (Date, Medical Specialty)	0 1 2	0 1 2	0 1 2
Day Hospital (Date, Medical Specialty)	0 1 2	0 1 2	0 1 2
Emergency Data (date, Medical Specialty, Cause, Destination after discharge)	0 1 2	0 1 2	0 1 2

Type of Data		Emergency			Inpatient			Outpatient		
CMD's										
Laboratories										
	Hematology	0	1	2	0	1	2	0	1	2
	Hemostases	0	1	2	0	1	2	0	1	2
	Immunology	0	1	2	0	1	2	0	1	2
	Pathologic anatomy	0	1	2	0	1	2	0	1	2
	Microbiology	0	1	2	0	1	2	0	1	2
Imagiology										
	X-Ray	0	1	2	0	1	2	0	1	2
	x-Ray Reports	0	1	2	0	1	2	0	1	2
	Echography	0	1	2	0	1	2	0	1	2
	Echography report	0	1	2	0	1	2	0	1	2
	CT	0	1	2	0	1	2	0	1	2
	CT report	0	1	2	0	1	2	0	1	2
	Magnetic Resonance	0	1	2	0	1	2	0	1	2
	Magnetic Resonance Report	0	1	2	0	1	2	0	1	2
	Angiography	0	1	2	0	1	2	0	1	2
	Angiography Report	0	1	2	0	1	2	0	1	2
	Osteodensitometry	0	1	2	0	1	2	0	1	2
	Osteodensitometry Report	0	1	2	0	1	2	0	1	2
Other CMD's (ECGs, Functional Tests, Endoscopy,...)		0	1	2	0	1	2	0	1	2
Other Documents										
Discharge Notes		0	1	2	0	1	2	0	1	2
Nursing Notes		0	1	2	0	1	2	0	1	2
Surgical Notes		0	1	2	0	1	2	0	1	2
Patient Managed Data (Pregnancy bulletin, Baby Bulletin, Vaccinations)		0	1	2	0	1	2	0	1	2

You reached the end of this questionnaire. **Thank you.**

COLOPHON

This document was typeset using a slightly adapted typographical look-and-feel `classicthesis` developed by André Miede. `classicthesis` is available for both LaTeX and LyX:

`http://code.google.com/p/classicthesis/`