

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO



FEUP

Information Systems Architecture definition for the Electronic Health Record (EHR)

Eduardo Filipe Garcia Pinto

Master in Informatics and Computing Engineering

Supervisor: António Carvalho Brito (PhD)

10th February, 2012

Information Systems Architecture definition for the Electronic Health Record (EHR)

Eduardo Filipe Garcia Pinto

Master in Informatics and Computing Engineering

Abstract

The Information Technologies have revolutionized many areas of society. Since many years ago, the health area has been searching for solutions to improve existing processes, having as the main objective the improvement of healthcare services. In this sense, the availability of patient clinical data can be vital to a more effective diagnosis and treatment, by a health professional. This information should be accessible regardless of context, place, time or where it was collected. In order to share this type of data, many countries have initiated projects implementing Electronic Health Record (EHR), including Portugal in 2010.

The EHR complexity and scope makes it an high cost and risk project. In fact, the inherent difficulties of a project like this are as critical as its importance and the impact it has on societies. The implementation of a successful EHR implies, in most cases, the involvement of all stakeholders, from patient to health organizations, along with the professionals themselves, becoming a very long and continuous process.

In the context of this dissertation we pretend to build a architecture proposal for implementing the Portuguese EHR. The proposed architecture should serve as a platform for sharing information between multiple health entities, normalizing processes, applying conventions and chasing interoperability among agents. The ability of the system evolution on one hand, and ease of integration in international projects on the other, are two major concerns to be taken into account throughout the proposal definition process.

This document reflects the study in order to understand what processes and practices are used in implementing these large-scale projects. Accordingly, we present some models of enterprise architectures as well as some architectural styles, both corresponding to high levels of abstraction. Regarding the integration of different systems, this document addresses several international conventions for transmission, encoding and storing of clinical information, such as HL7, openEHR and SNOMED CT. Also, we briefly analyse two International EHR case studies, in UK and Canada. In conclusion it should be noted that regardless of the technical challenges that may arise, the process must always be guided by the objective of improving healthcare services to the patient.

Keywords: Electronic Health Record; Health Information Systems; Interoperability; Information Systems Architecture

Resumo

As Tecnologias de Informação têm revolucionado as mais diversas áreas da sociedade. Desde há muitos anos atrás, a área da saúde tem procurado soluções tecnológicas que permitam melhorar os processos existentes, tendo como principal objetivo a melhoria dos cuidados de saúde que são prestados. Neste sentido, a disponibilidade da informação clínica sobre um paciente pode ser vital para um diagnóstico e tratamento mais eficazes, por parte de um profissional de saúde. Esta informação deve estar acessível independentemente do contexto, local ou data onde foi coletada. Com o objetivo de providenciar este tipo de dados, muitos países iniciaram projectos de implementação de Registo de Saúde Electrónico (RSE), entre os quais se inclui Portugal, em 2010.

A complexidade e abrangência faz do RSE um projecto de altos custo e risco. De facto, as dificuldades inerentes a um projecto deste género fazem juz à sua importância e ao impacto que tem nas sociedades. A implementação com sucesso de um RSE é um processo longo e implica, na maioria dos casos, o envolvimento de todos os agentes intervenientes no processo, desde o paciente à instituição de saúde, passando pelos próprios profissionais.

No contexto desta dissertação pretende-se construir uma proposta de arquitectura para implementação do RSE em Portugal. A arquitectura proposta deve servir como uma plataforma de partilha de informação entre as mais diversas entidades de saúde, normalizando os processos e codificações nesta área e promovendo a interoperabilidade entre os agentes. A capacidade de evolução do sistema por um lado, e a facilidade de integração em projectos internacionais por outro, são duas grandes preocupações a ter em conta ao longo do processo de definição da proposta.

O presente documento reflecte o estudo feito no sentido de perceber quais os processos e práticas existentes na implementação destes projectos de larga escala. Nesse sentido, apresentam-se alguns modelos de arquitecturas empresariais assim como alguns estilos de arquitectura, ambos correspondentes a níveis de elevada abstracção. No que diz respeito à integração de diferentes sistemas, são abordadas diversas convenções internacionais para transmissão, codificação e armazenamento da informação clínica, como o HL7, SNOMED CT ou openEHR. Ainda no âmbito deste documento, são analisados dois casos de estudo internacionais do RSE, no Reino Unido e no Canadá. Para concluir, importa sublinhar que, independentemente dos desafios técnicos que se possam apresentar, o processo deve ser sempre guiado pelo objectivo principal de melhorar os cuidados de saúde ao paciente.

Palavras-chave: Registo de Saúde Eletrónico; Sistema de Informação da Saúde; Interoperabilidade; Arquitectura de Sistemas de Informação

“Architecture is a verb, not a noun”

Christ J. Kamages

Contents

1	Introduction	1
1.1	Background	1
1.2	Context	2
1.3	Motivation and Objectives	2
1.4	Document Structure	3
2	Information Systems Models and Frameworks	5
2.1	Enterprise Architecture Frameworks	5
2.1.1	The Zachman Framework	6
2.1.2	The Open Group Architecture Framework	7
2.1.3	Federal Enterprise Architecture Framework	9
2.2	Architectural Styles	10
2.2.1	The Metropolis Model	10
2.2.2	Service-oriented Architectures	12
2.2.3	Event-driven Architectures	13
3	International Practices and Conventions	15
3.1	Health International Standards	15
3.1.1	Health Level Seven International	15
3.1.2	The openEHR	20
3.1.3	Digital Imaging and Communications in Medicine	22
3.1.4	Terminology and Ontology Standards	22
3.2	International Case Studies	23
3.2.1	Canada Health Infoway	24
3.2.2	National Health Service Connecting for Health	26
3.3	European Patients - Smart open Services (epSOS)	29
3.3.1	Services and Standards	30
3.3.2	Security Aspects	31
3.3.3	Identification and Authentication	31
3.3.4	Semantic Issues	31
3.3.5	System Architecture	32
4	Architectural Proposal	35
4.1	General architecture	35
4.2	Work plan	36
5	Conclusions	39

CONTENTS

References

41

List of Figures

2.1	Zachman Framework classification model	7
3.1	Approximate real-world usage of HL7 messaging standards	16
3.2	Example of an HL7 v2 acknowledgement message	17
3.3	Example of an HL7 v2 acknowledgement message	18
3.4	The ADL definition of Complete Blood Count archetype	21
3.5	EHRs Blueprint Architecture overview	25
3.6	The England's National Programme for IT	27
3.7	The NHS ITK Maturity Model	28
3.8	Basic Architecture of epSOS	33
4.1	Dissertation work plan	37

LIST OF FIGURES

Abbreviations

ADL	Archetype Definition Language
ADM	Architecture Development Method
ANSI	American National Standards Institute
CBSS	Community-Based Service Systems
CDA	Clinical Document Architecture
DICOM	Digital Imaging and Communications in Medicine
EDA	Event-driven Architecture
EHR	Electronic Health Record
EMR	Electronic Medical Record
FEAF	Federal Enterprise Architecture Framework
FEUP	Faculty of Engineering of University of Porto
HIS	Health Information Systems
HL7	Health Level 7
ICD	International Classification of Diseases and Related Health Problems
IHE	Integrating Healthcare Enterprise
IHTSDO	International Health Terminology Standards Development Organisation
IS	Information Systems
ISO	International Standardisation Organization
IT	Information Technology
LOINC	Logical Observations Identifiers Names and Codes
NHS	National Health Service
OSS	Open-source Systems
RIM	Reference Information Model
SNOMED	Systematized Nomenclature of Medicine
SOA	Service-oriented Architecture
SOAP	Simple Object Access Protocol
TOGAF	The Open Group Architecture Framework
WHO	World Health Organization
XML	Extensible Markup Language
ZF	Zachman Framework

ABBREVIATIONS

Chapter 1

Introduction

The revolution and advent of technology has been changing the way people interact with each other, forcing the companies to adopt new business strategies, improving life's quality and so forth. In the organizations context, the Information Systems brought a new opportunity for improving and optimizing processes, helping to generate value and profits and making them more competitive [GW91].

The health context is one of the most complex and critical, compelling the organizations to support several business processes and interacting with multiple stakeholders from clinicians to laboratory technicians. With such a complex scenario it seems easy to understand the impact that the technology might have, automating processes or sharing information among all the interested parties. However, this complexity obviously raises the risks of failing for projects implementation, either because the products do not meet the requirements, the stakeholders expectations or were built excessively monolithic what preclude future integration and evolution [Chu06].

1.1 Background

Despite of the healthcare arena largeness and the inherent difficulties of comprehensive Information Technology (IT) projects, several concepts and platforms started appearing [Hau06]. For instance, seems reasonable to say that the patient clinical information should be available to its consumers when needed, whatever the place or time of the occurrence. The Electronic Health Record (EHR) concept is described [GT05] as the “longitudinal collection of electronic health information about individual patients and populations”. The main objective is to provide clinical information about a patient where it needs to be consulted, independently of its origin or location, helping to avoid clinical errors or duplication of efforts and resources. It is supposed to be a mechanism for

integrating healthcare information for the purpose of improving care quality [Ors08] but obviously, these EHR systems are as complex as the implementations challenges they face.

1.2 Context

The benefits of an EHR seem consensual and undeniable. All around the world, several countries started expensive programmes trying to implement partial or total EHR solutions, some being more successful than others.

The Portuguese health information systems suffer from the same problems [Del11] that probably most of the countries: all over the years, the systems were being created essentially to solve local problems. These systems were developed without any kind of concern about interoperability — designing and building the systems following standards and not precluding future communications and data sharing. The consequence is quite predictable: the multiple existing systems, which have many purposes, use different codification standards and several interfaces to allow communication, do not have the capability to ‘talk’ with each other. Plus, even if they could communicate, they would not be able to understand each other since they do not use the same language to express themselves. As final consequence, there is an huge amount of clinical data that cannot be shared.

In 2010, the Ministry of Health signed a protocol with the Faculty of Engineering of University of Porto (FEUP) establishing a partnership for delivering some recommendations about the design and implementation for the EHR. This dissertation appears under the scope of that protocol.

1.3 Motivation and Objectives

Clements, Kazman and Klein advocate [CKK01] that “architectures allow or preclude nearly all of the system’s quality attributes”. In this sense, when it is pretended to create a system where the quality attributes are critical, the architecture is certainly one of the most relevant means to assure and facilitate them.

The purpose of this dissertation is to build an architecture proposal to the Portuguese EHR. This architecture must allow the integration of the existing multiple Health Information Systems, as well as providing an infrastructure for sharing data. The proposal must promote the interoperability between different healthcare organizations, either by forcing the utilization of international standards and conventions or by establishing well-defined interfaces. The future evolution of the system must also be taken into account when designing the architecture proposal.

Despite the deliverable is an architecture proposal, the final objectives can be pointed out as follows:

- deliver a full architecture proposal, able to offer a complete and integrated vision from the healthcare organization level to the EHR level;
- study the impact on the applications architecture describing the possibilities of new services integration;
- propose a governance model promoting the long-term management and evolution of the EHR;

The thesis research will be done through a straight collaboration with the Ministry of Health, allowing to work close to the EHR stakeholders.

1.4 Document Structure

In the present chapter we described the context, motivation and goals of this dissertation. The document has more 4 chapters.

The Chapter 2 starts with a small introduction about the HIS and try to understand what kind of methodologies may be used to facilitate the implementation of such large systems, using Enterprise Architectures. Then, we briefly describe three Architectural Styles.

In the Chapter 3 we start describing the implementation of two EHR case studies, in United Kingdom and in Canada. Next, we overview some of the health international standards and finish with a description of the epSOS programme.

The Chapter 4 aims to trace in general lines how will be structured the proposal and which are its main ‘components’ and concerns. Also, it is included the work plan for the next semester.

To conclude, the final Chapter 5 allows us to state some final considerations about the future work and which problems might raise up along.

Introduction

Chapter 2

Information Systems Models and Frameworks

In this chapter we present a general study about how is it possible to build ultra-large systems. The Information Systems (IS) are a kind of large dimension systems. The development methods used to implement these systems are an important source of experiences and techniques relevant to this dissertation. The first section describes some Enterprise Architecture Frameworks and in the second one we introduce two Architectural Models.

2.1 Enterprise Architecture Frameworks

The concept of Enterprise Architecture appeared more than 20 years ago. At that time, the systems complexity was growing with an exponential velocity. However, most of the times, those systems were not able to fulfil the business needs and the problem was not lack of technology or knowledge but difficulties in understanding the business from those who were developing it. Thus, the software development was facing two problems at a time: in one hand, the systems were becoming huge and hugely complex; on the other hand, the systems were developed with few concerns about business orientation [[Ses07](#)].

Despite of being a problem with a considerable age and very studied also, it did not stop growing and there is not a quite clear solution. A lot of enterprise architecture models appeared and disappeared over the years. However, it is estimated that around 90% of the market is dominated by this four methodologies [[Ses07](#)]:

- The Zachman Framework;
- The Open Group Architecture Framework (TOGAF);
- The Federal Enterprise Architecture Framework (FEAF);

In the next subsections, we will slightly describe each one of them, trying to understand which the advantages and weak points of each one.

2.1.1 The Zachman Framework

The Zachman Framework [Zac87] aims to guarantee that all stakeholders perspectives are being taken into account when developing a complex software system. In general terms, it is important to understand if all the artefacts are sufficiently focused and if the existing artefacts clarify all the players, from the business owner till the database designer, keeping all the visions aligned.

In the original article and in order to explain his methodology, John Zachman uses an analogy between building an information system and construct a building. Thus, when constructing a building, the architect starts with understanding the general requirements as how many divisions it is supposed to have and the general purpose of the building, usually by using bubble charts. In a second phase, the architect's drawings aims to represent the links between each room of the building and also its general structure. Finally, the architect presents the plans and they must represent and reflect the customer requirements and expectations. These plans already include several details about electrical system, masonry, wood structure, etc. After plans arrive the contractor, they suffer some adjustments either because of costs/price or because another kind of constructing limitations.

In general, we can say that there are three fundamental architectural representations, one for each player in this process: the owner, the designer and the builder. Through all this process, the project will be documented in many perspectives. The architect starts listening to the customer and building his own perspective of the project that is sent to the builder, which in turn builds another view that allows himself to clarify some aspects. Despite of all these documents refer and describe the same project, each one of them has different purposes. Each one of them is unique and independent, describes one area, one perspective, not being possible from one of those to conclude about the others.

Originally, the proposal of Zachman was based on a policy of three questions: What?, How? and Where?, used to describe the product. However, the model has evolved and now it is based in three more questions: Who?, When? and Why?.

Just as it is possible to see in Figure 2.1, Zachman claims a classification of the existing artefacts from multiple points of view, identifying:

- **Material description** — the data model of the system: “entity - relationship - entity”;
- **Functional description** — the process model of the system: “input - process - output”;
- **Location description** — the network model of the system: “node - line - node”;

	DATA <i>What</i>	FUNCTION <i>How</i>	NETWORK <i>Where</i>	PEOPLE <i>Who</i>	TIME <i>When</i>	MOTIVATION <i>Why</i>
Objective/Scope (contextual) <i>Role: Planner</i>	List of things important in the business	List of Business Processes	List of Business Locations	List of important Organizations	List of Events	List of Business Goal & Strategies
Enterprise Model (conceptual) <i>Role: Owner</i>	Conceptual Data/ Object Model	Business Process Model	Business Logistics System	Work Flow Model	Master Schedule	Business Plan
System Model (logical) <i>Role: Designer</i>	Logical Data Model	System Architecture Model	Distributed Systems Architecture	Human Interface Architecture	Processing Structure	Business Rule Model
Technology Model (physical) <i>Role: Builder</i>	Physical Data/Class Model	Technology Design Model	Technology Architecture	Presentation Architecture	Control Structure	Rule Design
Detailed Representation (out of context) <i>Role: Programmer</i>	Data Definition	Program	Network Architecture	Security Architecture	Timing Definition	Rule Speculation
Functioning Enterprise <i>Role: User</i>	Usable Data	Working Function	Usable Network	Functioning Organization	Implemented Schedule	Working Strategy

Figure 2.1: Zachman Framework classification model¹

- **People description** — the organization units and roles within the system: “organization - reporting - organization”;
- **Time description** — the events of the system: “event - cycle - event”;
- **Motivation description** — the high level organization goals of the system: “node - line - node”.

Over the years, this model started to be considered more like a taxonomy model, helping to frame all the artefacts and state their value and meaning to the multiple stakeholders [Ses07].

2.1.2 The Open Group Architecture Framework

First developed in 1995, The Open Group Architecture Framework was based on the US Department of Defense Technical Architecture Framework for Information Management (TAFIM). From this sound foundation, The Open Group Architecture Forum has developed successive versions of TOGAF at regular intervals and published them on The Open Group public web site [Jos11].

¹Source: http://upload.wikimedia.org/wikipedia/commons/d/da/Zachman_Framework_Detailed.jpg

TOGAF might be seen as a process for building an Enterprise Architecture. This framework states this building process as a continuous process of building multiple architectures from highly generic to highly specific ones, until reaching the organizational architecture level [[Ses07](#)].

This framework splits the EA in four categories:

- **Business Architecture** — describes processes used by business to meet own goals;
- **Application Architecture** — describes how to build applications from the framework and how to interact with the other ones;
- **Data Architecture** — describes how the information is stored, organized and accessed;
- **Technical Architecture** — describes the hardware and software supporting the applications and their interactions.

One of the most important concepts of this framework is the Architecture Development Method (ADM), which stands as a reliable and proven approach for building enterprise architecture descriptions, keeping the progress close to the business specific needs. Actually, it provides an overall process template for architecture development activity and a narrative of each architecture phase, describing each one in terms of objectives, approach, inputs, steps and outputs to the architects [[LZ06](#)].

The TOGAF ADM also provides several guidelines and techniques:

- **Architecture Content Framework** — detailed model for architectural work products, including deliverables and artefacts;
- **Enterprise Continuum** — a model for structuring a virtual repository and classify architecture and solution artefacts, tracking how the artefacts evolve and how they can be re-used;
- **TOGAF Reference Models** — two reference models: Technical Reference Model (TRM) and Integrated Information Infrastructure Model (III-RM);
- **Architecture Capability Framework** — a set of templates, guidelines, resources that aim to help the architect establishing some practices within an organization.

In overall terms, TOGAF ADM recommends an iterative process for architecture development, not being prescriptive on breadth of coverage, level of details or time horizons. The architect determines those details to fit one specific project. One of the most important characteristics of TOGAF ADM is that, despite of having well-defined phases, it enough flexible to be adapted to the organization and to let architects build the best approach to implement TOGAF [[TH04](#)].

2.1.3 Federal Enterprise Architecture Framework

The Federal Enterprise Architecture Framework appeared with the objective of serving as a platform for sharing processes, information and documentation among the U.S. Federal Agencies and other government agencies. This framework gathers two main characteristics of the two previous: in on hand, it defines a taxonomy – similar to the Zachman Framework (Section 2.1.1) – for artefacts classification; on the other hand, it suggests a process for building and implementing the architecture like TOGAF (Section 2.1.2) does.

The FEAF partitions a given ‘platform’ into business, data, applications and technologies architectures, aiming to take those into account when applying the process of implementing the new architecture. Furthermore, the framework is organized in four levels [TH04]:

- Level I — the highest level which deals with the architecture drivers or external stimulus fetching the strategic direction of the architecture. It facilitates the transformation from the original architecture to the new one by applying architecture standards and managing the process;
- Level II — an analyses of the business goals, direction, principles, strategies and priorities;
- Level III — expression of architecture in more detailed view by using business data, applications and technology views to model it;
- Level IV — representation of Data Architecture, Application Architecture and Technology Architecture using a combination of Zachman Framework and Spewak’s Enterprise Architecture Planning methods.

FEAF also suggests a measurement process, being able to evaluate the maturity the inner organizations are adopting and implementing the enterprise architecture. This evaluation classifies the agencies measuring three essential attributes:

- Architectural completion — the maturity of the architecture itself;
- Architectural use — how is the architecture being used to improve efficiency of decision-making;
- Architectural results — the benefits brought by the use of the architecture.

Based on the analysis of this three attributes, the agencies are then rated in one of three categories: green, yellow and red.

2.2 Architectural Styles

An Architectural Style is a particular pattern that organizes components and connectors in a particular form, providing a well-known reliable solution when facing a specific kind of problem [Kaz]. Each architectural style principle influences some quality attributes in a positive and some other in a negative way. Zheng Qin *et al* define it as “a solution to solve a certain class of problems which have common quality attributes requirements”, stating also that “there is no architecture style that is proper for all systems, because every system have different quality attributes requirements” [QZX08]. Kim and Garlan advocate [KG10] that these concepts bring a number of significant benefits as they promote design reuse.

2.2.1 The Metropolis Model

Over the last years, two important trends are changing business and society: in one hand, the growing use of shared knowledge to create high-value services and projects; on the other hand, the growing preponderance of services. The business are evolving from a product-oriented perspective to a relation-oriented one. This change of paradigm brought the client to the middle of the business, helping to create value. The greatest examples of this new philosophy are the systems based in communities as Wikipedia or Facebook, that are able to generate profit and value directly from its users.

The Metropolis Model [KC09] appears as an attempt of describing really huge complex systems built from two basilar concepts: Open-source Software (OSS) and Community-Based Service Systems (CBSS). Despite the growing importance of OSS and CBSS, there will always be some systems that cannot be developed by crowds, either because of the business criticality or for demanding unquestionable security guarantees. The Metropolis Model does not really apply to those systems.

We will start with an explanation of the ‘crowdsourced systems’ and then briefly describe the principles of the Metropolis Model, always following the original article wrote by Rick Kazman and Hong-Mei Chen [KC09].

2.2.1.1 Crowdsourced Systems

The Metropolis Model identify its target as the ‘crowdsourced systems’. This concept includes the kind of systems with some special characteristics:

- open teams — meaning that the assumption of having a closed and dedicated team of programmers should be abandoned;
- ‘mashability’ — as the capacity of composing and integrating different systems and functionalities since that the consumption of software by one person or project does not make it less available for consumption by another;

- conflicting, not knowable requirements — the requirements emerge from the contributors and the users, making them highly unpredictable and often being in conflict between each other;
- continuous evolution — as a consequence of constant changing requirements and distributed resources, these type of systems are never finished nor stable, introducing the ‘perpetual beta’ concept. The development of the project is done through multiple iterations empowering the users with the responsibility of testing and assuring its quality;
- focus on operations — the availability of its operations are, most of the times, as critical as their utility and, in that sense, it is not accepted any kind of system downtime (Amazon, eBay or Google are good examples);
- sufficient correctness — the notion of ‘perpetual beta’ expresses the idea of acceptance some incompleteness in software (e.g. Wikipedia);
- unstable resources — one of the most interesting facts is that, despite of the volatility of people, information and other resources that individually would be useless, when massive congregated are able to create reliable, stable and impressive computational powered systems;
- emergent behaviours — frequently, there are behaviours and tendencies that emerge from the system that are completely unexpected. In these systems, unlike the traditional ones, the crowds take the rudder and bring to the system unpredicted kinds of utilization or interaction.

2.2.1.2 The principles

The Metropolis Model presents a new unified vision between the CBSS and the OSS, focusing deliberately in the crowd value generation. This model suggests the creation of two levels: the kernel services and the periphery services. It states that the application’s kernel may be developed using traditional methods.

The Metropolis Model is supported by the application of some principles:

- **Crowd engagement and egalitarian management of open teams** — the management must focus the people and the crowds, engaging the users for value co-creation. The engagement is not just a system-level issue, but a question of business strategy. These kind of system resort to crowd-sourcing chasing “potential for cost reduction, increased innovation, and quicker development time for delivering products and services that meet customer needs” as the authors explain;

- **Bifurcated requirements** — the requirements must be distinguished into kernel or periphery ones. The kernel should “deliver little or no end-user value” (e.g. Linux kernel, Wikipedia wiki or Facebook application platform). On the other hand, from the periphery appears most of the end-user value, as it happens with Wikipedia articles or Facebook applications;
- **Bifurcated architecture** — the architecture of these systems are composed by the kernel infrastructure and the peripheral services. Since the kernel must be designed and implemented to be stable and promote the integration of several components, it must not emerge from the use of the system;
- **Fragmented implementation** — the crowd-sourced philosophy must be applied only to the peripheral services. The kernel implementation shall be done by an “close-knit, highly motivated, coordinated team”, ensuring its high quality.
- **Distributed testing** — while the kernel must be highly reliable and constantly tested, in the peripheral services it is only necessary the sufficient correctness;
- **Distributed delivery/maintenance** — at the kernel level, it is important to preserve backwards compatibility when deploying a new version. At periphery the release mechanisms are uncoordinated and form a constant stream;
- **Ubiquitous operations** — these systems are “always on” even when they are being upgraded. Also, they should monitor self state and create control mechanisms not allowing peripheral problems to affect the system’s core.

2.2.2 Service-oriented Architectures

A Service-oriented Architecture (SOA) is “an architectural style that emphasizes implementation of components as modular services that can be discovered and used by clients” and that “emphasis on loose coupling between interacting services” [ST05]. However, as Thomas Erl stated [Erl05] there is no single definition of SOA. Instead, there are many opinions about what constitutes service-orientation. Erl advocates that the service-orientation paradigm has its “roots in a software engineering theory known as separation of concerns”. Obviously, this concept has been applied in many contexts and to many platforms, from object-oriented programming till component-based programming approaches.

The service-orientation paradigm has no official principles. Although, there are “a common set of principles most associated with service-orientation” [Erl05]:

- **Services are reusable** — the services are designed to promote and support their reuse, regardless of whether immediate reuse opportunities exist;

- **Services share a formal contract** — in order to be possible to share, there is a necessity for establish the terms of information exchange;
- **Services are loosely coupled** — services must be designed to avoid dependency outside itself;
- **Services abstract underlying logic** — the only part visible to the outside is the interface and the requirements that need to be met to obtain the service, and that are clarified in the contract's description;
- **Services are 'composable'** — services might be aggregated with the finality of creating larger services, obtaining different abstraction levels and promoting re-utilization;
- **Services are autonomous** — the service is governed within an explicit boundary not being dependent to perform its operations;
- **Services are stateless** — services should not be required to manage state information as that might preclude their ability to stay independent and not coupled. The concerning about maximizing the statelessness must be central.

To conclude, as Srinivasan and Treadwell advocates [ST05] “service-orientation reinforces general software architecture principles such as encapsulation, modularization and separation of the interfaces from their implementations”.

2.2.3 Event-driven Architectures

The Event-driven Architectures (EDA) are based in the event concept. An event [Mic06] is a notable thing that starts inside or outside the system and may consist in a problem, an opportunity and so forth. The event meaning should be defined in the business context. In an EDA, when an event is triggered it is immediately disseminated to every interested parties [QZX08]. Then, these parties evaluate the event and might either call a service, trigger another component or simply take no action.

This architectural pattern promote loosely coupled and highly distributed systems. In fact, the creator of the event only has the task of dispatching it. From that point on, the creator loses track of the event, not knowing nothing about the post processing or even who are the interested parties. Usually, these architectures are defined by three main concepts:

- **event trigger** — triggers the event sending it to the event collector;
- **event collector** — redirects the received events to the interested parties;

- interested party — receives the event information and proceeds in accordance, calling another component or service;

This basic structure might be scaled, allowing to organize and build much complex systems.

To conclude, it is important to say that these architectures are advisable to be used for asynchronous flows of work and information. Also, these architectural styles easily combine for creating robust solutions [[Mar06](#), [LC08](#)].

Chapter 3

International Practices and Conventions

In this chapter, the objective is to present the main Health International Standards for interoperability, helping us to understand which options there are to allow the systems to properly share, interpret and store data. In addition, we analyse two EHR implementations case studies, from Canada and United Kingdom. Finally, we present a small overview about the epSOS project, trying to evidence the most important architecture factors that might influence the Portuguese EHR.

3.1 Health International Standards

The pursuit of interoperability is not possible without the clearly definition of common languages and communication channels. These common languages are called standards and are usually created and managed by independent organizations.

3.1.1 Health Level Seven International

Health Level Seven International (HL7) is one of several accredited organizations of American National Standards Institute (ANSI), operating in the healthcare area. HL7 provides a set of standards towards interoperability, aiming to facilitate knowledge transfer among several stakeholders: healthcare providers, government agencies, patients and so forth [[Sev](#)].

3.1.1.1 HL7 Messaging Standard

One of the most used standards is the messaging one. The HL7 Messaging Standard appeared several years ago and had been evolving along the last two decades. Several institutions adopted and implemented it all over the world, most of the times investing lots of

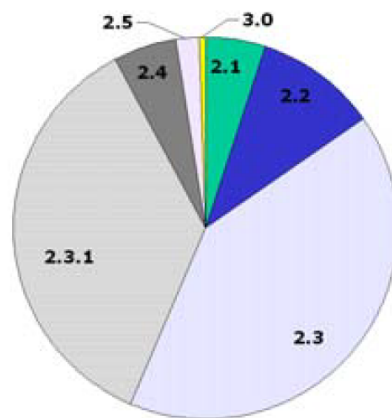


Figure 3.1: Approximate real-world usage of HL7 messaging standards [Sha10]

time and money to achieve it. Dave Shaver stated some interesting numbers [Sha10] about the utilization of the different standard versions, which are resumed in Figure 3.1. By observing the Figure 3.1, we can easily conclude that “the vast majority of HL7 messaging is done using messages that approximate HL7 2.3 or HL7 2.3.1” unlike the newer ones (2.5 or greater and 3.0) that represent a “very small portion of real-world interfaces” [Sha10].

3.1.1.2 HL7 Messaging Standard Version 2

The HL7 v2 is a standard created several years ago. Nevertheless, it had evolving into some new versions. Currently, HL7 v2 is at version 2.7. However, as it was designed as backwards compatible, the version is not so relevant. In fact, what is often more important is how the system which is sending the message is populating the fields and segments the standard defines.

The HL7 v2 stands as a simple protocol for exchanging clinical data and it stands as the most widely used healthcare information standard [EAR⁺05]. Certainly, one reason for that is that the referred messages are plain text, with several fields split by the vertical slash character ‘|’ as it is possible to observe in Figure 3.2.

However, some authors advocate that “the messages lack a formal underlying reference model, and feature considerable optionality in addition to permitting user-defined elements”. In fact, that excessive flexibility obliges to establish a prior agreement structure and interpretation that might be called “negotiated interoperability” [AKPW10]. Obviously, this vagueness “provides great flexibility but necessitates detailed bilateral agreements among the healthcare systems to achieve interoperability” as state Eichelberg *et al* [EAR⁺05].

In this context, it is important to clearly identify the main advantages of using HL7 v2, which we point out next [AKPW10]:

¹Source: <http://www.interfaceware.com/images/ACKMessageAnatomy.png>

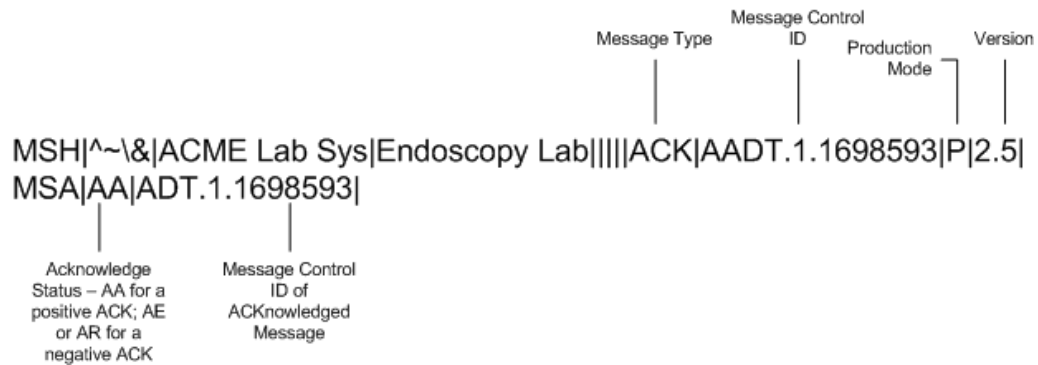


Figure 3.2: Example of an HL7 v2 acknowledgement message¹

- richly expressive data types;
- support for arbitrary structured objects;
- easy to adopt and use with low cost.

As expected, there are some arguments against also, which are:

- excessive flexibility and optionally in message interpretation implies a “negotiated interoperability”;
- lack of privacy and consent guarantees;
- growing tendency to abandon v2 in favour of v3.

To conclude, despite of being an exceeded standard it is still very relevant, being used world wide.

3.1.1.3 HL7 Messaging Standard Version 3

The HL7 v3 appeared as a natural evolution of HL7 v2, trying to solve some problems that the previous version had. Its first version was launched in 2005. Despite of being just a new release, this update means significant standard modifications as the main objective was to “establish semantic interoperability in loosely coupled systems” [AKPW10]. In this sense, at the core of this standard just appeared the Reference Information Model (RIM).

The RIM is an object model created as part of the Version 3 methodology, serving as a context provider to the multiple healthcare concepts. For instance, a WBC (white blood count) may refer to an order (intent) or a result (observation). In that case, the RIM structures “implement the context by binding the coded vocabulary term from the ontology to a specific place in the model” [Sha04].

```

<identifiedPerson>
  <!-- Primary id as used/known by this registry -->
  <id extension="000197245" root="2.16.840.1.113883.19.3"/>
  <addr use="H">
    <streetName>Randomroad</streetName>
    <houseNumber>25a</houseNumber>
    <postalCode>1200</postalCode>
    <city>Anytown</city>
  </addr>
  <telecom use="H" value="tel:555 3542557"/>
  <statusCode code="active"/>
  <identifiedPerson>
    <name use="L">
      <given>Patricia</given>
      <family>Patient</family>
    </name>
    <administrativeGenderCode code="F" codeSystem="2.16.840.1.113883.5.1"/>
    <birthTime value="19750103"/>
    <maritalStatusCode code="F" codeSystem="2.16.840.1.113883.5.2"/>
  </identifiedPerson>
  <assigningOrganization>
    <!-- Scoper, the registering organisation -->
    <id extension="1002777" root="2.16.840.1.113883.19.200"/>
    <contactParty nullFlavor="UNK"/>
  </assigningOrganization>
</identifiedPerson>

```

Figure 3.3: Example of an HL7 v3 message [Spr07]

With HL7 v3, the messages are no longer plain text but XML files (as we can see in Figure 3.3) which turns the message more human-readable but also more structured and rigid.

There are some advantages that should be credited to this Version 3, which are important to summarize [AKPW10, Sha10]:

- rich in clinical semantics;
- address many of the practical difficulties such as incomplete information, uncertainty, duplicate records and so forth;
- dynamic model no more complex than v2;
- “less expensive to build and maintain mid-to-long term interfaces”;
- “more of a ‘true standard’ and less ‘framework for negotiation’”;
- the long journey of maturing and reflecting about the standard.

Notwithstanding of the several advantages, there are some drawbacks that can be identified [AKPW10, Sha10]:

- not clear where the return of investment is since it does not offers substantial advantages over v2 for some areas (simple alerting, drug interaction checking, recall systems, best practice guidelines, clinical pathways, and so forth);

- message structures complex and difficult to understand;
- “inconsistencies between an Information Model (objects document things) interpretation and a Reference Ontology (objects are things) interpretation”;
- not suitable as a model for storage EHR information and does not specify and EHR Architecture neither;
- unclear how is it possible to query a v3 messages repository;
- absence of compatibility with HL7 v2;
- expensive and slow adoption.

In an interesting research by Gartner in 2006, Rishel predicted that “the HL7 V3 messages will fail to achieve critical mass, being replaced by ever-more-elaborate use of V2, a series of dialects of V3 chosen by individual countries and large enterprises, or the work of other standards organizations that may achieve less semantic interoperability while being easier to implement” [Ris06].

To complete, the HL7 v3 seems the natural evolution for the applications using the previous versions. Although, that migration will take many years and it might not be so pacific as would be desirable.

3.1.1.4 Clinical Document Architecture

Clinical Document Architecture (CDA) is an ANSI-certified standard from HL7 organization. The first version (Release 1.0) was released in November, 2000 and the second (Release 2.0) was published with the HL7 2005 Normative Edition. [Sev] CDA defines the semantic and structure of clinical documents aiming to allow its exchanging. The CDA documents are actually XML documents. That concept is derived from the fact that these documents “derive their meaning from the HL7 Reference Information Model (RIM) and use the HL7 Version 3 Data Types which are part of the HL7 RIM” [AKPW10].

The CDA document is constituted by two parts, an header and a body. The header is used to unambiguously state the semantics of each entry in the document. The body “contains the clinical document content and can be either an unstructured text, or comprised of nested containers” [EAR⁺05].

It is important to distinguish the two versions of this standard, since there are many applications that still use the first version. The main characteristic of CDA standard is the documents classification into three different levels, although the interpretation of them is differentiated between the two versions.

The CDA Release One considers two levels [Sev, EAR⁺05]:

- Level One — basically it has a standardized header and additional information on the rest of the document;

- Level Two — allows to define and constraint both the structure and the content of the document, favouring the interoperability between systems since the receiver knows what to interpret;

The approximation of CDA Release Two is a little different. In fact, it provides an incremental approach, allowing and encouraging the documents to evolve and becoming more structured. One of the major additions is that the CDA body “uses RIM structures and controlled vocabulary to permit a much higher level of semantic interoperability” [Sev]. Atalag *et al* advocate that it is a “stable and flexible standard for developing communications with coded information” as well as “provides a pathway from narrative free text documents through incrementally coded data to fully coded and semantically interoperable clinical data” [AKPW10].

3.1.2 The openEHR

The openEHR standard is a standard owned by the openEHR Foundation, created in 2002. The Foundation is a non-for-profit organization with the founding partners being University College London and Ocean Informatics, an Australian company. The main objective is to research and develop interoperable electronic health records related topics, aiming to improve the care quality to patients [Les07].

Unlike other standards, openEHR develops specifications for implementing full EHR systems, pronouncing more in persistence as opposed to messaging. openEHR states that, in order to achieve lifelong, patient centred, secure and shareable EHR, the way cannot be the aggregation of messages [Bea02].

In order to provide an high level of interoperability, openEHR suggests the engaging of standardized clinical content models (called Archetypes), a stable reference model and semantically rich terminologies [AKPW10]. More in detail, Bale advocated a two-level methodology to model the EHR structure. In the first level, the existence of a generic reference model, specific to the healthcare domain but sufficiently general to be stable over the time (e.g. role, act, entity and so forth). In the second level, there were concepts mapped as archetypes, such as blood pressure, laboratory results and so on. The archetypes apply “constraint rules that specialize the generic data structures that can be implemented using the reference model” [Bea02].

Beyond the reference information model, the openEHR framework includes other resources that help the implementation following this standard, such as the ADL (Archetype Definition Language) language for expressing archetypes, an archetype library and a collection of open source implementations. As an example (Figure 3.4), we can restrict a generic Observation class to, for example, a Blood Pressure archetype.

```
OBSERVATION[at1000.1] matches {-- complete blood picture
  name matches {
    CODED_TEXT matches {
      code matches {[ac0001]} -- complete blood count}}
  data matches {
    LIST_S[at1001] matches {-- battery
      items cardinality matches {0..*} \epsilon {
        ELEMENT[at1002.1] matches {-- haemoglobin
          name matches {
            CODED_TEXT matches {
              code matches {[ac0003]} -- haemoglobin}}
          value matches {
            QUANTITY matches {
              value matches {0..1000}
              units matches {^g/l|g/dl|.+.}}}}
        ELEMENT[at1002.2] occurrences matches {0..1} matches
        {-- haematocrit
          name matches {
            CODED_TEXT matches {
              code matches {[ac0004]}-- haematocrit}}
          value matches {
            QUANTITY matches {
              value matches {0..100}
              units matches {"%"}}}
        ELEMENT[at1002.3] occurrences matches {0..1} matches
        {-- platelet count
          name matches {
            CODED_TEXT matches {
              code matches {[ac0005]} -- platelet count}}
          value matches {
            QUANTITY matches {
              value matches {0..100000}
              units matches {"/cm^3"}
            }}}}}}
```

Figure 3.4: The ADL definition of Complete Blood Count archetype [EAR⁺05]

3.1.2.1 Advantages and Disadvantages

There are some interesting characteristics about openEHR that might be useful to underline [AKPW10]:

- intuitive and understandable model for clinicians;
- approach based on recording and querying observations;
- extracts can be sent with HL7 v2 (see 3.1.1.2);
- likely stable reference model over the time;
- possibility to convert archetypes into CDA documents (see 3.1.1.4).

On the other hand, Atalag *et al* identified some arguments against either [AKPW10]:

- “lacks semantic rigour and does not contain a logically sound ontology”;
- works well with simple scenarios, but does not easily handle complexity;
- no experiences with medium to large-scale systems;

- weak governance of the openEHR Foundation.

To conclude, the interest with this standard is spreading worldwide. There is an increasing number of full or partial implementations of the openEHR specifications in several countries, like United Kingdom, Sweden, Australia, Denmark, The Netherlands, Singapore, USA, Japan, Brazil, Scotland or Turkey. However, there is not a really large-scale systems in neither of these countries and the use of the standard consists of using some of its concepts [AKPW10]. In another interesting article, Marta Silva and José Carvalho state that the “openEHR standard has original and sound fundamental concepts that definitely will influence future generations of health information systems” as long as note that “some openEHR core beliefs are highly debatable”. Also, they raise some questions about the applicability of semantic interoperability “when the industry all over the world is struggling in exchanging basic unstructured demographic and clinical data through the complex network of health providers”. [SC11]

3.1.3 Digital Imaging and Communications in Medicine

DICOM (Digital Imaging and Communications in Medicine) is an worldwide used standard for medical image communication. The standard provides data structures and services allowing the exchange of medical images and related information. In the actual structure, the standard is available since 1993, despite the creation remounts to 1983.

Unlike most of other EHR standards, the DICOM uses a binary encoding. Pianykh has an interesting point of view, saying that “contrary to popular belief, DICOM is not just an image or file format” but “an all-encompassing data transfer, storage, and display protocol built and designed to cover all functional aspects of digital medical imaging” [Pia08].

Other authors stated that “it has become a leading standard used by all major vendors of diagnostic medical equipment”, predicting also that “DICOM will soon be used in every medical branch that utilizes imaging, for example: cardiology, mammography, radiology, surgery, endoscopy, dentistry, pathology, etc” [MDG08].

3.1.4 Terminology and Ontology Standards

The need of data sharing between different healthcare institutions leverage the creation of multiple standards, attempting to allow, not only the data sharing, but also the easy interpretation of the message’s content. In fact, this kind of internationally endorsed classifications facilitate the storage, retrieval analysis and interpretation of data. In this sense, in the next subsections we will present some terminologies that were created with the aim of making the systems understand each other.

3.1.4.1 Systematized Nomenclature of Medicine - Clinical Terms

SNOMED CT (Systematized Nomenclature of Medicine - Clinical Terms) was created in 2006 by the International Health Terminology Standards Development Organisation (IHTSDO). This standard aims to provide a unique and embracing system with clinical terms. The objective is to have a repository, managed and updated centrally, available to all systems which adopted it and that can be used either to clinical purposes as to research projects. [[ACS09](#)]

Citing the official site, the SNOMED CT “provides the core general terminology for the electronic health record (EHR) and contains more than 311,000 active concepts with unique meanings and formal logic-based definitions organized into hierarchies” and “can be used to represent clinically relevant information consistently, reliably and comprehensively as an integral part of producing electronic health records” [[IHT](#)].

3.1.4.2 International Classification of Diseases

The International Classification of Diseases (ICD) [[WHO](#)] is “the international standard diagnostic classification for all general epidemiological, many health management purposes and clinical use”. This standard is owned by the World Health Organization (WHO) and dates from 1850, being developed and reviewed every ten years. Although, every year new updated are release to the versions in use. The ICD provides an huge variety of codes to classify diseases, body signals, symptoms, abnormal aspects, complaints, social circumstances and external causes of injury or illness, beyond to having an additional classification used to classification of transplants and newborns and so forth.

ICD is a classification, and a essential one, as it define the universe of entities to be studied, and highlight the relevant aspects of the information that has been collected. Also, it allows their comparison in several contexts: within and between populations over time and the compilation of internationally consistent data.

3.2 International Case Studies

Several Electronic Health Record projects were initiated in multiple countries. The study of same of those might be fundamental in order to understand what can we learn with them. By doing that, it will help us to extend our horizons, retain the bad experiences and better evaluate some options that we might have to take.

In the next subsections, we will provide an overview by two EHR initiatives: the Canada Health Infoway and the National Health Service (from England).

3.2.1 Canada Health Infoway

Canada Health Infoway is a non-for-profit organization founded by Canada's First Ministers in 2001. It was specifically created to accelerate the process of development of Electronic Health Record systems, promoting the adoption of standards that guide to communication facilitation between different healthcare organizations.

In order to guide the development of the systems in each different province, Infoway provided a national framework called EHR Blueprint. The EHR Blueprint is a set of principles, guides and components. It states "a comprehensive description of the components necessary for the interoperable EHR and describes, in broad terms, how the components are envisioned to work together" [Inf06].

3.2.1.1 Sharing EHR Information

In the process of building an EHR, there are several methods to allow sharing EHR information along several services, consumers and providers. EHRs Blueprint advocates that the best method (at least for their reality) is the creation of a shared reference information source that is populated by several health-care organizations around Canada. This reference is populated with clinical relevant data and is maintained externally from every health-care organization (or Points of Service, as designated in EHRs Blueprint). The Points of Service (PoS) are able to reference or pull data from the shared repository.

The 'EHR Infostructure' is based on achieving full integration and interoperability between the EHR Solution and each PoS. In order to do so, EHRs Blueprint has a Health Information Access Layer that provides the interface which all the PoS communicate with. However, since EHRs Blueprint defines a unique interface for all PoS, most of these systems needed to adapt themselves to respect the standards and be able to stay connected to the system.

This shared reference presupposes the existence of peers along each different Canada's jurisdiction. These peers represent several copies of the EHRi in terms of structure, but dealing only with the local PoS. As the infrastructure is the same, the process of retrieving information from other peers, when needed, is relatively simple.

3.2.1.2 Architectural principles

The EHRs Blueprint states [Inf06] some architectural principles which guided the EHRs implementation. However, we will just point out the ones that might be more relevant and interesting at this point.

The fact of the EHR Infostructure information being stored as copy of the original one is a key characteristic, as it preserves independence between the EHRi and the PoS.

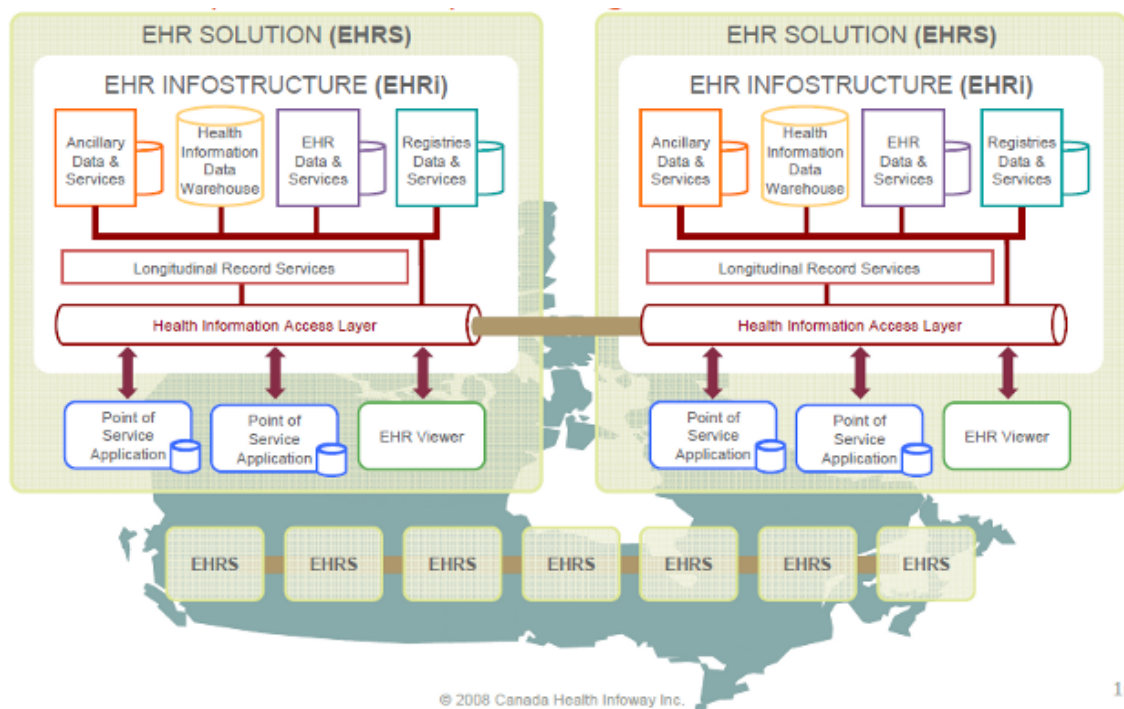


Figure 3.5: EHRs Blueprint Architecture overview [Inf]

Another relevant principle is the controlled environment built around the system, defining one common interface and transforming EHRi into a black-box in which PoS can retrieve but also update clinical information for a specific patient.

The EHRs Blueprint was built following an Services Oriented Architecture, making the architecture more flexible and the components reusable.

Another interesting fact is that there is no single ‘home’ for the patient’s electronic health record. Actually, each jurisdiction’s EHRi holds and owns the data generated in health services from that jurisdiction.

3.2.1.3 Key elements

The EHR Blueprint has multiple components with certain characteristics. Although some of those had been already referred, it is useful to take a deeper look at them and point out some others also. Thus, the key elements are [Inf06]:

- Point of Service Applications (PoS)** — these are software applications or information systems that provide clinical information to the EHR system, working as gateway for gathering critical patient data, absolutely fundamental to the system’s purposes. It is important to notice that these systems are responsible for most collection of the patient’s EHR data. Here, we refer to applications as an information system in hospital emergency department as well as an local pharmacy system and further so;

- **EHR Data Repositories** — sometimes, there is relevant clinical information that is not available in the PoS applications despite of being very relevant in a clinical decision making context. Instead, it is usually available through other systems. In this sense, the PoS applications are also responsible for pushing the data into these EHR Data Repositories – that become responsible for storing it and keeping it available to the users that might need that. Four logical clinical domain repositories are identified by EHRS Blueprint: Shared Health Record, Drug Information, Diagnostic Imaging and Laboratory;
- **Registry Services** — there are the information linkers. These Registry Services provide the identification of patients, matching them with the required clinical data. In order to guarantee the match of required and retrieved information, these services offer: Client Registry, Provider Registry, Location Registry and Terminology Registry;
- **Longitudinal Record Services (LRS)** — as we saw in Subsection 3.2.1.1, the EHRS Blueprint is based on distributed data repositories. Thus, when it is needed to retrieve and show the information to the user (for instance, a physician), all the data must be gathered as if it was stored in the same place. These Longitudinal Record Services execute that task, bringing together data from different registries and sources, normalizing it for common understanding;
- **Health Information Access Layer (HIAL)** — it provides a single standardized way of sharing and retrieving data from EHRi. The fact of being unique and the single entry point obligates the PoS applications to adapt themselves to the interface but also to the information standards also.

3.2.2 National Health Service Connecting for Health

The National Health Service (NHS) Connecting for Health is part of the UK Department of Health. This organization was created on 1 April 2005 as the replacement for an older one (NHS Information Authority). It has the responsibility of managing and implementing the NHS National Programme for IT (NPfIT). The NPfIT was an initiative to upgrade the NHS to a centrally-mandated electronic health record for patients, connecting hundreds of hospitals and thousands of healthcare professionals and providing them relevant patient data by secure and certified means.

3.2.2.1 Architecture overview

The NPfIT has been born as the world's largest civil information technology project, committing £12.4 billion over 10 years in order to improve the quality of healthcare in England.

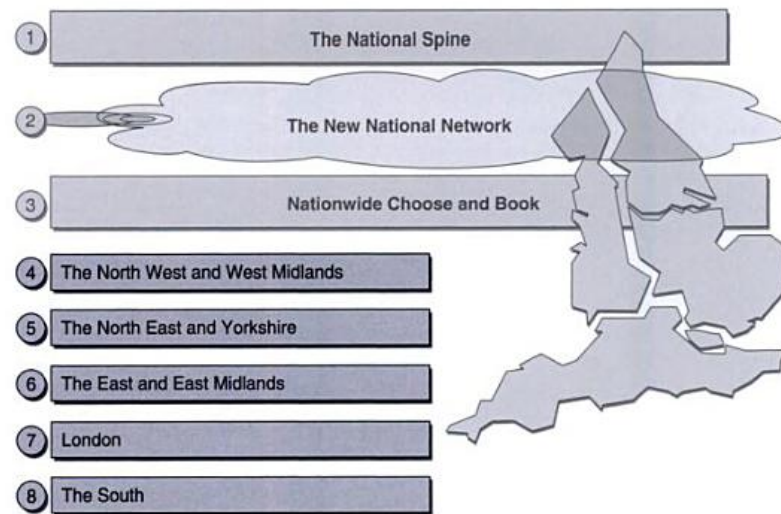


Figure 3.6: The England's National Programme for IT [Bre05]

As expected, such a big project could not be just a single standalone implementation. In fact, the NPfIT is made of eight separated systems, which are [Bre05]:

- **One National Application Service Provider (NASP)** — designed to support the 'National Data Spine', which was destined to keep the patients' electronic data;
- **A New National Network (N3)** — a network connecting all the hospitals, applications and other healthcare providers, characterized by the need of being a really broadband one. It would be an essential infrastructure supporting and linking all the other services providers;
- **One Electronic Appointment Booking (eB)** — a service (now called 'Choose and Book') that would give the possibility of the patients to choose the place, date and time for their appointments in a hospital or clinic;
- **Five Local Service Provider (LSP)** — five physically distributed providers, hosting the NHS Care Records Service (NHS-CRS) at a local level, covering all England's territory.

The Figure 3.6 briefly describes the kind of interaction between the different components of the system. It is important to notice that, in order to implement the five local clusters, five providers were contracted and made responsible for delivering the local services. The idea was that, in one hand, the providers would be challenged to compete between each other, speeding up the process of implementation. On the other hand, the risk would be lower since there was different suppliers implementing similar systems in parallel. CSC Alliance, BT Health London, Accenture and The Fujitsu Alliance were the contracted LSP's for the main body of the programme.

3.2.2.2 NHS Interoperability Toolkit (ITK)

The NHS Interoperability Toolkit is a set of standards, frameworks and implementation guides to support and favour the interoperability between local systems and across them. The NHS ITK aims to support flexibility and local innovation and also removing barriers to entry. It also wants to be an enabler of evolution and reusing of solutions that already proved to be valuable, connecting all the peers by standardization in order to ensure that there are no silos being created.

One of the key concepts of ITK is the use of a maturity-based approach, allowing the organizations to evolve through small steps, particularly for CDA documents. The step-by-step maturity model allows the organization to incrementally progress from sharing binary data to sharing fully-coded CDA documents.

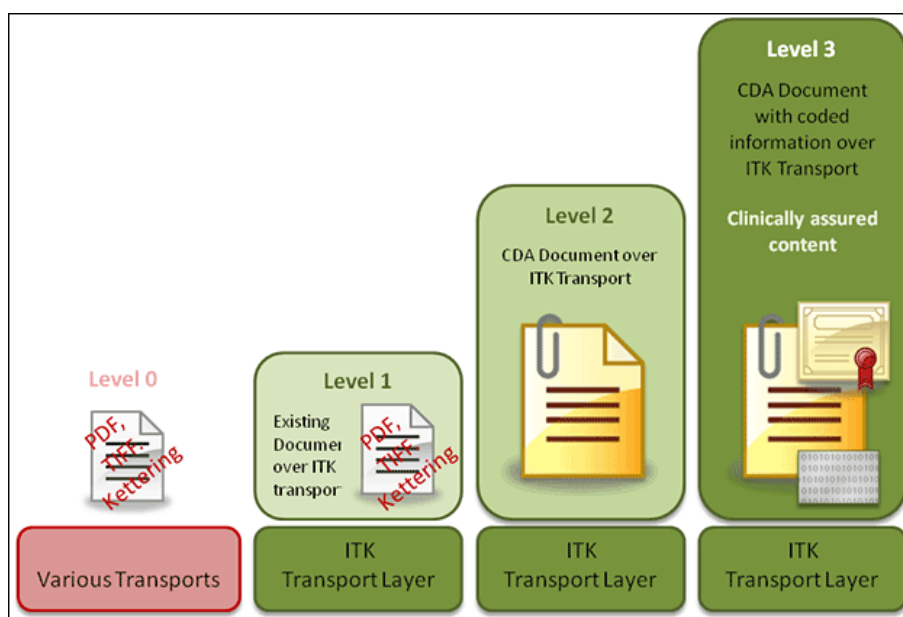


Figure 3.7: The NHS ITK Maturity Model [fH12]

As it is possible to observe in Figure 3.7, there are four possible levels in the model. The lowest one – Level 0 – is applied to any organization, when even the sharing itself is made out of the ITK Transport Layer. When one organization becomes to use the ITK Transport Layer it reaches the Level 1. Then, the Level 3 is assigned when an organization is able to share CDA documents. Finally, when the data is passed through CDA documents one organization has reached the highest level – Level 4. In Section 3.1.1.4 we explain the CDA standard more in detail.

3.2.2.3 NHS Care Records

The NHS Care Records is one of the NPfIT's components. This component is what we usually call an Electronic Health Record system, aiming to provide personal clinical information to the healthcare providers, increasing the quality and efficiency of the treatments.

The NHS Care Records considers two different types of records:

- **Summary Care Records** — records held nationally. A Summary Care Record – usually called Patient Summary – stores essential information about a person, available in emergency situations, informing the health professionals about what medicines are one taking, the allergies that might suffer from or any known bad reactions to other medicines;
- **Detailed Care Records** — records held locally. The Detailed Care Record is a more comprehensive record which might store data from past exams and details, avoiding the necessity for repeating them, for example.

3.2.2.4 The fall and the failure

The NPfIT started in October 2002 and since then it always been the target of some criticisms. However, in April 2006, a set of 23 academics, wrote an open letter² raising several questions and concerns about the programme. A report by the King's Fund in 2007 also criticised the government's "apparent reluctance to audit and evaluate the programme", questioning their failure to develop a capable strategy [WAH07].

The next years were very troubled as long as several reports raising doubts about the feasibility of the project were made public [Pow04, Coi07, Bre07, CS07, Bre09]. One of those, by the Public Accounts Committee, stated in 2009 that the risks for the programme deployment were "as serious as ever", bringing up serious uncertainty about the capacity of the systems to meet expectations of clinical staff.³ The over-and-over delays to deliver the fundamental services and the overrun of the programme's cost started to being unsustainable. In September 2011, the NPfIT has been dismantled following the conclusions of a new review by the Cabinet Office's Major Projects Authority (MPA) [oH].

3.3 European Patients - Smart open Services (epSOS)

The epSOS (European Patients - Smart open Services) is the main European electronic Health (eHealth) project created by European Commission and the partners. The main objective is to improve the health assistance to citizens who are abroad their own country.

²Details: http://editthis.info/nhs_it_info/The_Open_Letter_to_the_Health_Select_Committee

³Details: <http://news.bbc.co.uk/2/hi/health/7850619.stm>

It aims to do so by implementing an interoperability platform, allowing different countries to exchange vital information. In that context, one health professional could access patient data from another country improving and précising the medical treatment [EpS].

3.3.1 Services and Standards

The epSOS is based on a Services-oriented Architecture. It states that all services are passive, implemented with traditional Web Services with their interfaces based on the Web Service Description Language [W3C WSDL 1.1]. When we say that the services are passive, it means that every transaction must be initiated by the consumer of the information. The key element of epSOS is the National Contact Point (NCP) and it stands as the only peer of contact for one determined country. In this sense, an NCP can serve both as a content provider when the request is made from other countries or as a gateway for an information request from that country, conveying that request to the target one.

The NCP is based on a set of Common Components and when connecting those components to the national infrastructure it is possible to offer the following end-user services: Consent Service, Patient and Order Services, eDispensation Service, Auditing Service.

3.3.1.1 Consent Service

epSOS states that the patient must consent and allow the sharing of his information between peers of the epSOS network. This consent is realized through the IHE Basic Patient Privacy Consents (BPPC) profile⁴ and the patient's consent is given in the origin country.

3.3.1.2 Patient and Order Services

The Patient and Order Services are used to retrieve medical information when abroad of the patient's home country. The medical data is returned in the epSOS common format and is realized by the IHE Profile XCA⁵ Cross Gateway Query and Cross Gateway Retrieve. These documents are transmitted by web services exchanges and are complaints with the HL7 CDA Release 2.0 standard (see 3.1.1.4).

There are two type of documents to be exchanged: the 'ePrescription' and the 'Patient Summary', both using epSOS specific CDA Level 1 and Level 3, respectively.

⁴Details: http://wiki.ihe.net/index.php?title=Basic_Patient_Privacy_Consents

⁵Details: http://wiki.ihe.net/index.php?title=Cross-Community_Access

3.3.1.3 eDispensation Service

The eDispensation Service is required whenever it is necessary to update some medical data with some new information. In this context, one ‘eDispensation’ document is dispatched from the actual country to the patient’s home country, being formatted with HL7 CDA Level 3.

3.3.1.4 Auditing Services

The Auditing Services aim to guarantee that all that sensible information is only accessible by those who have that right. Thus, it defines some rules about how to achieve Auditing and Authentication using the IHE Profile Audit Trail and Node Authentication⁶.

3.3.2 Security Aspects

In terms of security, epSOS states some basic non-functional security requirements in order to guarantee that all transactions and data are reliable. Thus, epSOS must ensure: Identification, Authentication, Access Control, Non-repudiation, Data confidentiality, Data availability and Logging activities that impact security. Beyond these generic requirements, epSOS establishes some specific security requirements to be used in each country, dividing those into three levels:

- First level — general epSOS as a whole (16 requirements);
- Second level — general NCP (34 requirements);
- Third level — national information infrastructure (9 requirements).

3.3.3 Identification and Authentication

3.3.4 Semantic Issues

epSOS is obviously an huge and complex project, presenting several challenges to manage and resolve. We know that each country has its own policy, classification tables, ontologies and so on. So, one of the most interesting challenges is how to handle this diversity, finding a common language which enables the communication and transcription of several local classifications and taxonomies, not to mention the country language itself. In order to address this problem, epSOS defines a set of semantic services which we describe next.

⁶Details: http://wiki.ihe.net/index.php?title=Audit_Trail_and_Node_Authentication

3.3.4.1 CDA's PPC in epSOS

The CDA's PPC (Patient Care Coordination⁷) was integrated in epSOS, creating a significant amount of data elements.

3.3.4.2 epSOS Master Value Sets Catalogue

The epSOS Master Value Sets Catalogue (MVC) contains all value sets used within epSOS system and CDA framework. The value sets were built based on several well-known code systems, like SNOMED CT, ICD-9, ICD-10, LOINC, ATC, HL7 and so forth. The epSOS MVC is distributed as an Excel file.

3.3.4.3 epSOS Master Translation/Transcoding Catalogue

The epSOS Master Translation/Transcoding Catalogue (MTC) is based on the epSOS MVC. It is made to fit the specific needs of a country and it is managed and updated its representatives. It is also provided as an Excel file.

3.3.4.4 epSOS Ontology

The epSOS Ontology aims to unify different terminologies that might be used in the future, providing a linguistic reference of the terms in the epSOS value sets.

3.3.5 System Architecture

epSOS architecture (described in Figure 3.8) was built with the objective of being enough flexible to allow integration with all kind of services existing inside the boundaries of each member state. The Figure 3.8 presents the peer-to-peer connection between two countries. Also, it is possible to understand which components stay under the member state responsibility.

We will overview the main functions of each component in the next subsections, following a top-down approach.

3.3.5.1 epSOS Interface

The epSOS Interface is one of the common components, representing the higher layer of the system. This component might assume one of two roles: Income Protocol Terminator (IPT) when acting as NCP-A or Outbound Protocol Terminator (OPT) when acting as NCP-B.

When Member State B (the consumer of some required information) dispatches a request, the OPT transforms the objects into SOAP requests at the same time that signs

⁷Details: http://wiki.ihe.net/index.php?title=Patient_Care_Coordination

Basic Architecture

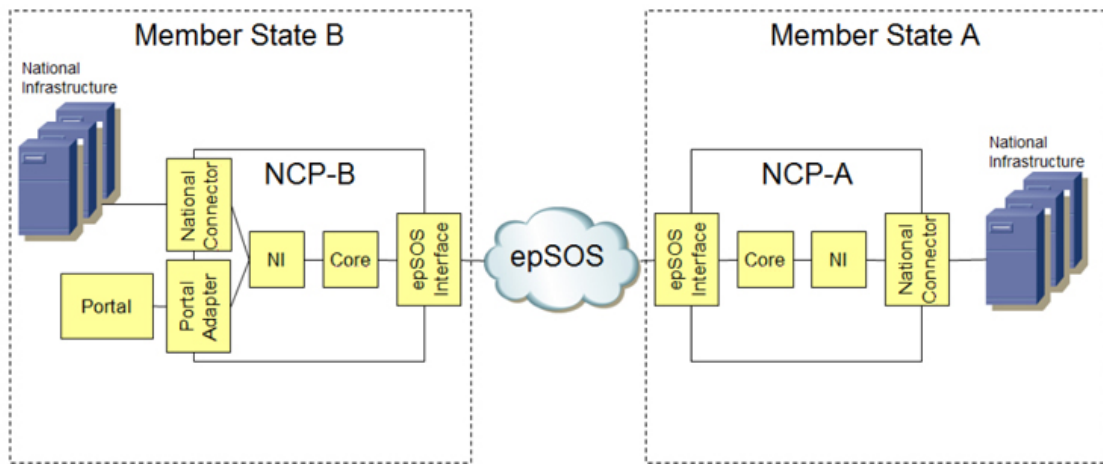


Figure 3.8: Basic Architecture of epSOS [EpS]

and certifies that request and send it to Member State A (the target NCP). At this point, the request arrives at the ICP of NCP-A. The ICP verifies the authenticity of the message and, if reliable, transcode the SOAP request into objects and convey them to a lower main component called Workflow Manager. On the other side, the NCP-B waits a response and when it arrives, it does exactly the same, returning the information to the its respective Workflow Manager. The Workflow Manager is one of the core components and will be exposed later.

3.3.5.2 Core elements

The core elements serve as the business layer of the system. These are the components which are considered core elements:

- **Workflow Manager** — it is the entry point for the business layer of NCP, working like an orchestrator and being called both by IPT and the National Connector. The Workflow Manager is the controller of all the operations process, being able to call other services available by other components interfaces, usually returning that information to the National Connector to the OPT;
- **Security Manager** — it is the responsible for certifying and validating documents;
- **Transformation Manager** — it works as a data translator from the local language to the epSOS Reference Terminology and vice-versa;
- **Terminology Services Access Manager** — called by Transformation Manager. It is responsible for translating a local concept into a epSOS coded concept one, using

the Terminology Repository, which represents the epSOS Reference Terminology, and is managed and updated by the country;

- **Audit Trail Writer** — it is the responsible for creating an event log of every transaction and sending it to the Audit Repository. Although, the audits are limited in terms of information.
- **Audit Repository** — it is the responsible element for storing the logs produced above. The logs are only accessible to the national infrastructure. Thus, these audit interfaces might be defined taking into account the national specific needs;
- **Routing Manager** — it stands as the pointer to the other countries, providing the routes to connect them. the routing tables are XML documents that might be stored locally or be fetched from a central repository either.

3.3.5.3 The National Interface and Connector

The National Interface represents the connection between the Workflow Manager and the National Connector. The National Connector is a collection of adapters that provide access and make use of the internal nation services, using the local protocols and formats. As it is the link for the national information systems, it can also invoke the Workflow Manager through the National Interface.

Chapter 4

Architectural Proposal

This dissertation pretends to deliver an Architectural Proposal for the EHR. In that sense, this chapter is a briefly overview about how the work will be structured and the fundamental questions that must be answered in the end of this dissertation project.

4.1 General architecture

At this point, there are three essential questions that resume the work to be done:

- How to align the technology and business visions, allowing future evolution of the system?
- Which data will be stored where?
- Which standards will be used to format data and exchange information?

The problem of align the technology and business visions is one the greatest problems when talking about ultra-large scale software projects, as it was already referred in Section 2.1. The solution will probably come from the intersection of recommendations from the several Enterprise Architecture frameworks referred.

As the EHR is supposed to be a long-term project, the necessity for evolution will be a constant through the years. In that sense, it is critical that the architecture proposal allow and promote that evolution. Some approaches that might be helpful were described in Section 2.2.

The amount of information that is produced and stored, every day, by the healthcare organizations is enormous. This information can be classified by importance level what can give us an idea about which data is the most relevant and must be always available. Thus, the relevant information storing is a nuclear issue, as it can vary from a totally centrally-based approach to a complete distributed one.

To conclude, the way information is shared around the ‘players’ includes the definition of: how to transmit the data, how to interpret the data and how to store the data. Basically, the problem starts with putting the systems to share information. Then, it is also required that the information transmitted can be understandable by the receiver. Finally, the way the information is stored, either centrally or in each institution is critical to facilitate all the integration.

4.2 Work plan

A work plan was created to guide the dissertation research work and is expressed in Figure 4.1. The plan was divided in several tasks:

- International standards — the first weeks will be necessary to study some of the standards internationally adopted, helping to understand better the EHR needs in terms of information standards;
- Service-Oriented Architectures — research about the implementation of this type of services in the health context and its applicability to the EHR;
- Governance Model — some time dedicated to study how it will be possible to manage all the system, maintaining the data integrity while allowing evolution;
- Applications Architecture — the requirements that allow new systems to be integrated and share information with the existing ones;
- Architecture Proposal — the proposal will be created over the time, along with the research;
- Evaluate the Proposal — evaluation with the study and application of some well-known methods;
- Write and review thesis report — finally, some time reserved to writing and reviewing the thesis document.

Architectural Proposal

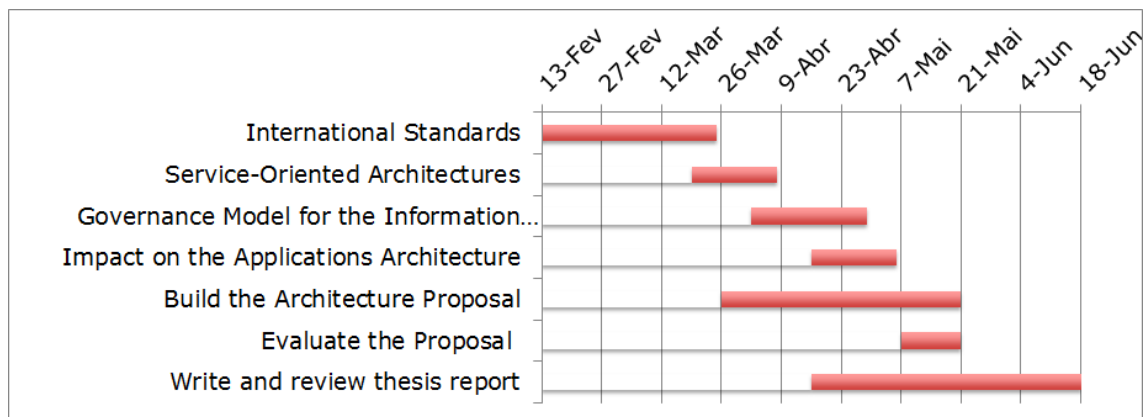


Figure 4.1: Dissertation work plan

Architectural Proposal

Chapter 5

Conclusions

The importance of an Electronic Health Record system and its benefits in the healthcare services equals the difficulty that they face to be implemented. In fact, these kind of platform take many years to be implemented, resulting in high costs. In this context, it is important that this architecture proposal can reflect the existing case studies, in order to avoid approaches that could led to implementation failures. However, the challenges are not only the size of such project. Most of the times, EHR implementations have to deal with and manage multiple external factors, either political, economical or social. These factors tend to limit and condition the execution progress.

At this phase, it is already possible to state some critical success factors in terms of the final architecture proposal. The health area has several actors, multiple institutions and not less business processes. In this scenario, it will be fundamental that the research for designing the proposal can be done close to the main stakeholders, essentially the healthcare organizations IT directors. On the other hand, the Portugal economic situation demands these projects to obtain short-term results. In this dissertation, the main objective is not the short-term perspective but the long-term evolution and vision of the system. However, it is also important that the architecture proposal allows to obtain some quick results, without precluding the future development, not only because of the economic situation but also because that might be the right way to engage all the EHR stakeholders, from patients to healthcare organizations.

The validation of the architecture might be an important point in order to increase confidence and decrease the implementation risks. Despite of that analysis might not be trivial, some evaluation software architecture techniques might be explored and adapted to this context. Due to the project specificities it was essential to close and clearly define the scope of this dissertation, under the penalty of becoming too sparse.

Apart from the conceptual and business challenges, the final purpose must guide all

Conclusions

the future work: improving the patient healthcare services allowing the healthcare professionals to access relevant clinical information.

References

- [ACS09] ACSS/MS. Documento de Estado da Arte. Technical report, 2009.
- [AKPW10] Koray Atalag, Douglas Kingsford, Chris Paton, and Jim Warren. Putting Health Record Interoperability Standards to Work. *electronic Journal of Health Informatics*, 5(1):1–17, 2010.
- [Bea02] Thomas Beale. Archetypes: Constraint-based domain models for future-proof information systems. In *OOPSLA 2002 workshop on behavioural semantics*, pages 1–18. Citeseer, 2002.
- [Bre05] Sean Brennan. *The NHS IT project: the biggest computer programme in the world - ever!* Radcliffe Publishing Ltd; 1 edition (April 2005), 2005.
- [Bre07] Sean Brennan. The biggest computer programme in the world ever! How’s it going? *Journal of Information Technology*, 22(3):202–211, July 2007.
- [Bre09] Sean Brennan. The National Programme for IT (NPfIT): is there a better way? *Integrating Healthcare with Information and Communications Technology*, page 95, 2009.
- [Chu06] S C Chu. From component-based to service oriented software architecture for healthcare. *Proceedings of 7th International Workshop on Enterprise networking and Computing in Healthcare Industry 2005 HEALTHCOM 2005*, 4(1):96–100, 2006.
- [CKK01] Paul Clements, Rick Kazman, and Mark Klein. *Evaluating Software Architectures: Methods and Case Studies*. Addison-Wesley, Boston, MA, USA, 2001.
- [Coi07] Enrico W Coiera. Lessons from the NHS National Programme for IT. *The Medical journal of Australia*, 186(1):3–4, January 2007.
- [CS07] Chris Clegg and Craig Shepherd. ‘The biggest computer programme in the world...ever!’: time for a change in mindset? *Journal of Information Technology*, 22(3):212–221, July 2007.
- [Del11] Deloitte. Saúde em análise - Uma visão para o futuro. Technical report, 2011.
- [EAR⁺05] Marco Eichelberg, Thomas Aden, Jörg Riesmeier, Asuman Dogac, and G.B. Laleci. A survey and analysis of Electronic Healthcare Record standards. *ACM Computing Surveys (CSUR)*, 37(4):277–315, December 2005.

REFERENCES

- [EpS] EpSOS. epSOS: Systems & Standards. <http://www.epsos.eu/technical-background/systems-standards.html>.
- [Erl05] Thomas Erl. *Service-oriented architecture: concepts, technology, and design*. Prentice Hall PTR, 2005.
- [fH12] Connecting for Health. NHS Interoperability Toolkit - NHS Connecting for Health. <http://www.connectingforhealth.nhs.uk/systemsandservices/interop>, 2012.
- [GT05] T.D. Gunter and N.P. Terry. The emergence of national electronic health record architectures in the United States and Australia: models, costs, and questions. *Journal of Medical Internet Research*, 7(1), 2005.
- [GW91] Vijay Gurbaxani and Seungjin Whang. The impact of information systems on organizations and markets. *Communications of the ACM*, 34(1):59–73, 1991.
- [Hau06] Reinhold Haux. Health information systems - past, present, future. *International journal of medical informatics*, 75(3-4):268–81, 2006.
- [IHT] IHTSDO. IHTSDO: About SNOMED CT. <http://www.ihtsdo.org/snomed-ct/snomed-ct0/>.
- [Inf] Canada Health Infoway. Model Report. <https://knowledge.infoway-inforoute.ca/EHRSRA/index.html>.
- [Inf06] Canada Health Infoway. EHRS Blueprint v2 Executive Overview. (April):1–33, 2006.
- [Jos11] Andrew (The Open Group) Josey. TOGAF ® Version 9.1 Enterprise Edition. *Group*, pages 1–13, 2011.
- [Kaz] Rick Kazman. Software Architecture. *Architecture*, pages 1–21.
- [KC09] Rick Kazman and Hong-Mei Chen. The metropolis model a new logic for development of crowdsourced systems. *Communications of the ACM*, 52(7):76, July 2009.
- [KG10] Jung Soo Kim and David Garlan. Analyzing architectural styles. *Journal of Systems and Software*, 83(7):1216–1235, July 2010.
- [LC08] Zakir Laliwala and Sanjay Chaudhary. Event-driven service-oriented architecture. *Service Systems and Service*, 2008.
- [Les07] Heather Leslie. openEHR - The World’s Record. *Pulse+ IT Magazine*, (6):50–55, 2007.
- [LZ06] Susanne Leist and Gregor Zellner. Evaluation of current architecture frameworks. *Proceedings of the 2006 ACM symposium on Applied computing SAC 06*, page 1546, 2006.

REFERENCES

- [Mar06] Jean-Louis Maréchaux. Combining service-oriented architecture and event-driven architecture using an enterprise service bus. *IBM Developer Works*, (April):1–8, 2006.
- [MDG08] Mario Mustra, Kresimir Delac, and Mislav Grgic. Overview of the DICOM Standard. In *ELMAR, 2008. 50th International Symposium*, volume 1, pages 39–44. IEEE, 2008.
- [Mic06] BM Michelson. Event-driven architecture overview. *Patricia Seybold Group*, 2006.
- [oH] Department of Health. Dismantling the NHS National Programme for IT. <http://mediacentre.dh.gov.uk/2011/09/22/dismantling-the-nhs-national-programme-for-it/>.
- [Ors08] P.R. Orszag. Evidence on the Costs and Benefits of Health Information Technology. In *Testimony before Congress*, volume 24, 2008.
- [Pia08] O.S. Pianykh. *Digital imaging and communications in medicine (DICOM): a practical introduction and survival guide*, volume 3. Springer Verlag, 1 edition, 2008.
- [Pow04] J Powell. NHS national programme for information technology: changes must involve clinicians and show the value to patient care. *BMJ: British Medical Journal*, 328(7449):1200, May 2004.
- [QZX08] Zheng Qin, Xiang Zheng, and Jiankuan Xing. *Architectural Styles and Patterns*. 2008.
- [Ris06] Wes Rishel. HL7 V3 messages need a critical midcourse correction. *Gartner Industry Research Publication ID*, (G0014095):3–8, 2006.
- [SC11] Marta Silva and José Carvalho. openEHR utilization in real systems. *Information Systems*, 2011.
- [Ses07] Roger Sessions. A Comparison of the Top Four Enterprise-Architecture Methodologies, 2007.
- [Sev] Health Level Seven. HL7 - FAQs. <http://www.hl7.org/about/FAQs/index.cfm?ref=nav>.
- [Sha04] Mark Shafarman. HL7 Version 3: the use case for the Reference Information Model, 2004.
- [Sha10] Dave Shaver. The HL7 Evolution. 2010.
- [Spr07] René Spronk. HL7 Message examples: version 2 and version 3. http://www.ringholm.de/docs/04300_en.htm, 2007.
- [ST05] L. Srinivasan and J. Treadwell. An overview of service-oriented architecture, web services and grid computing. *HP Software Global Business Unit*, 2, 2005.

REFERENCES

- [TH04] Antony Tang and Jun Han. A comparative analysis of architecture frameworks. *Software Engineering Conference*, 2004.
- [WAH07] Derek Wanless, J Appleby, and A Harrison. Our future health secured. *A review of NHS funding*, 2007.
- [WHO] WHO. WHO | International Classification of Diseases (ICD). <http://www.who.int/classifications/icd/en/>.
- [Zac87] J. a. Zachman. A framework for information systems architecture. *IBM Systems Journal*, 26(3):276–292, 1987.