OraMod

VPH based predictive model for oral cancer reoccurrence in the clinical practice

TITLE	**D2.4 Technical standard position paper**		
Deliverable No.	D2.4		
EDITOR	ONETONET		
Contributors	A.Ruggeri, O.Notelli, G.Scoponi		
WorkPackage No.	WP2	**WorkPackage Title**	Usage specifications and system design
Status[1]	Draft	**Version No.**	3
Dissemination level	PU		
DOCUMENT ID	OraMod D2.4 Technical standard position paper		
FILE ID	OraMod_D2.4_draft_final_3.6.pdf		
Related documents	OraMod_D2 2_final.pdf – Usecases and user needs OraMod_D2 3_final.pdf – Technical design		

Distribution List

Organization	**Name of recipients**
UNIPR	T. Poli
VUMC	R. H. Brakenhoff, M van de Wiel
Fraunhofer	F. Jung, S. Steger, S. Wesarg
ST- Italy	Sarah Burgarella, Marco Cereda
VCI	G. Nikitas, G. Aristomenopoulos
OneToNet	A. Ruggeri, O. Notelli, S. Copelli
UDUS	K. Scheckenbach
VTT	M. Kulju
European Commission	EC Officers and experts

[1] Status values: TOC, DRAFT, FINAL

Revision History

Revision no.	Date of Issue	Author(s)	Brief Description of Change
1	20/04/2014	A. Ruggeri	First Draft
2	19/05/2014	A.Ruggeri, G.Scoponi	Final Draft
3	20/05/2014	E. Martinelli	Added Revisions and comments by Coordinator, to be upgraded
4	12/06/2014	A.Ruggeri, G.Scoponi	Final draft after revision

List of abbreviations and definitions

AJAX	Asynchronous JavaScript and XML
AJCC	American Joint Committee on Cancer
API	Application Programming Interface
ATAG	Authoring Tool Accessibility Guidelines
CDA	Clinical Document Architecture
CII	Inter-Institutional Committee for Informatics
CORBA	Common Object Request Broker Architecture
CRAN	Comprehensive R Archive Network
CSS	Cascading Style Sheets
CT Scan	Computed Tomography Scan (Diagnostic Imaging)
DB	Data Base
DOM	Document Object Model
DoW	Description of Work, Technical Annex I to the Grant Agreement
EC	European Commission
ECMA	European Computer Manufacturers Association
EHR	Electronic Health Record
EJB	Enterprise JavaBeans
EMA	European Medicine Agency *http://www.ema.europa.eu/*
EU	European Union
FDA	Food and Drug Administration
FOSS	Free and Open Source Software
GIS	Geographic Information Systems
GiST	Generalized Search Tree
GNU	*Recursive of* "GNU's Not Unix" (O.S. Unix like)
HIS	Hospital Information System
HL7	Health Level 7
HTML	HyperText Markup Language
IDA (Programme)	Interchange of Data between Administrations
IHE	Integrating the Healthcare Enterprise
IP Management	Intellectual Property (IP) Management

ISBN/ISSN	International Standard Book Number/International Standard Serial Number
JCP	Java Community Process
LIS	Laboratory Information System
LOINC	Logical Observation Identifiers Names and Codes
MBAN	Medical Body Area Networks
MVCC	Multi-Version Concurrency Control
ODBC	Open Database Connectivity
OpenFTS	Open Source Full Text Search engine
OSEPA	Open Source software usage by European Public Administrations
OS	Operating System
OSS	Open Source Software
PACS	Picture Archiving and Communication System
PHR	Personal Health Records
RDBMS	Relational Data Base Management System
RIM	Reference Information Model
RMI	Remote Method Invocation
RPM	Remote Patient Monitoring
SVG	Scalable Vector Graphics
TCO	Total Cost of Ownership
TESS Programme	Telematics in Social Security
TESTA	Trans-European Services for Telematics between Administrations
UAAG	User Agent Accessibility Guidelines
W3C	World Wide Web Consortium
WAI-ARIA	Web Accessibility Initiative – Accessible Rich Internet Applications
WCAG	Web Content Accessibility Guidelines
XML	Extensible Markup Language
XSL-FO	Extensible Stylesheet Language - Formatting Objects

Table of Contents

List of figures

List of tables

1 Introduction

This document indicates the technology standards to be adopted for system interoperability by OraMod.

It is structured in sections, where each one defines a set of standard to be adopted in the project.

The first section – this one – outlines how the document is structured and mainly identifies the generic aspects of the document itself. The second one depicts the reference environment – as is and to be -. The third section identifies all the standards referenced in the previous paragraph. Then a series of Appendixes follow, related to the different parts of the projects: the first one explains the IDA Programme Document Specifications, a Community programme concerning interchange of data between administrations; the second appendix relates the Open Source Software Approach; the following' is a panoramic over the Software Tools suggested in the project; the fourth illustrates the Web-based Distributed Architecture with Java environment and the W3C standards; in the fifth appendix there is a section dedicated to the Relational Data Base; the sixth one is dedicated to the IHE interoperability and HL7 standards; the following appendix is dedicated to the DICOM standards for imaging; the last treats about the Medical Software Security Standards.

1.1 Purpose

To fulfill the project requirements, and mainly to achieve the two main goals of this segment of the project (i.e. a distributed and interoperable platform in the clinical environment) we will particularly focus on architecture standards and on integration standards.

To realize this document the IT standards adopted by participating hospitals for their Hospital Information Systems (HIS) have been collected and a survey on most relevant standards used in eHealth has been conducted by OneToNet. In particular the standards related to web-based distributed architecture were considered, as well as main W3C standards regarding web browser interoperability, accessibility (W3C Web Accessibility Initiative - WAI) and internationalization (I18n Activity).

IHE Patterns and HL7 messaging system will be the key logical architecture for the data integration workflows. In particular we will use HL7 3.0 CDA (Clinical Document Architecture) rel. 2 standard (ISO standard, ISO/HL7 27932:2009) for clinical data interoperability For HL7 messaging and parsing we will adopt HAPI classes (HL7 2.x - 3.0 parser project for JAVA initiated by University Health Network in Toronto).

1.2 Scope

This document is intended to be a working guideline for OraMod partners, to allow them to use a common set of standards, certain and shared, just to perform similar, compatible developments and use same procedures.

The document would constitute a reference for the technological committees and address the EU Commission towards similar health projects and models.

The document is structured in order to constitute also a reference for all the partners by a technological point of view.

As such this internal deliverable may be used also by the health professionals in charge of quality assurance within the project.

The document therefore uses notations and terminology specific for the OraMod project operators.

1.3 References

This section list all the applicable and reference documents, identified by title, author and date. In alternative other documents present in the net or links to a define environment are directly addressed.

Num.	Title (Applicability & Reference)	Author	Date
[1]	Ares (2011)1016907	EU	September 26, 2011
[2]	Open Source software usage by European Public Administrations	INTERREG IVC	November 11, 2012
[3]	FDASIA Health IT Report	FDA/ONC/FCC	April 2014
[4]	Cybersecurity for Medical Devices and Hospital Networks	FDA	June 13, 2013

Num.	Title (Applicability & Reference)	Link
[5]	IDA Programme	http://ec.europa.eu/enterprise/ida/index.htm
[6]	Redmine	http://www.redmine.org/
[7]	Apache Subversion	http://subversion.apache.org/
[8]	Eclipse	http://www.eclipse.org/
[9]	R	http://www.r-project.org/
[10]	ORACLE - JAVA EE	http://www.oracle.com/technetwork/java/javaee/overview/index.html
[11]	W3C - JAVASCRIPT WEB APIs	http://www.w3.org/standards/webdesign/script
[12]	W3C - XML Technology	http://www.w3.org/standards/xml/
[13]	W3C - INTERNATIONALIZATION	http://www.w3.org/standards/webdesign/i18n
[14]	W3C - ACCESSIBILITY	http://www.w3.org/standards/webdesign/accessibility
[15]	IHE	http://www.ihe.net/Profiles/
[16]	HL7	http://www.hl7.org/implement/standards/index.cfm?ref=quicklinks
[17]	LOINC	http://en.wikipedia.org/wiki/LOINC http://loinc.org/
[18]	DICOM	http://medical.nema.org/standard.html

Table 1 - References

2 The reference environment

Here follows the current state of the art present in the three Hospitals of the project and how OraMod project will change it by a physical and logical point of view.

2.1 Clinical Partners Technical Survey

During the early stage of the project we performed a technical survey in order to assess the different systems and the area of technical competence of the IT structures in every single clinical center.

The first goal of this analysis was to define the interoperability standards to be used and their compatibility with existent HIS and Clinical Systems and the activity volumes of the specific wards involved in the project.

In addition we analyzed the specific IT technologies and tools expertise of the Clinical partners IT personnel in order to choose, in a later stage, the Interoperability Standards, the Operating Systems, the DBMS and the Software Tools on which they were already skilled and confident.

Last goal of this survey was to assess the Hardware architecture and resources available in every Clinical Partner structure before the start of the project and potential upgrade/update plans during the project lifecycle.

2.2 Tech survey results

	HNO	VUMC	UNIPR
Hospitalization Ward (Involved in the ORAMOD Project)			
Number of beds	45	34	18
Number of hospitalizations/year	about 2000	1574	About 1000
Average Hospital Stay (days)	4,7	4,27	4,5
Number of Medical Doctors	9	10,63	16
Number of Nurses	2	22,5	15
Outpatient Clinic (Involved in the ORAMOD Project)			
Number of outpatient rooms	6	16	
Number of outpatient visits/year	about 10000	26339	
Number of Medical Doctors	5	10,63	16
Number of Nurses	4	3,7	
HIS - Hospital Information System			
Demographic Registry SW - Name and Brand	Siemens Medico	PATREG	ADS-Gruppo Finmatica
Demographic Registry SW - HL7 Integration possible?	yes	Yes	yes
Demographic Registry SW - HL7 Version	ver 2.3	2.2, 2.4	
Demographic Registry SW - Non-HL7 Integration Patterns			
Patient registration (ADT) Software - Name and Brand	Siemens Medico	PATREG	AREAS-Engineering

	HNO	VUMC	UNIPR
Hospitalization Ward (Involved in the ORAMOD Project)			
Patient registration (ADT) Software - HL7 Integration possible?	yes	Yes	yes
Patient registration (ADT) Software - HL7 Version	ver 2.3	2.2, 2.4	
Patient registration (ADT) Software - Non-HL7 Integration Patterns		Webservice	
Order Scheduler - Name and Brand	Siemens Medico	UltraGenda	AREAS-Engineering
Order Scheduler - HL7 Integration possible?	yes	Yes	yes
Order Scheduler - HL7 Version	ver 2.3	2.4/2.5	
Order Scheduler - Non-HL7 Integration Patterns		Webservices	
EPR (Electronic Patient Record) - Name and Brand	Siemens Medico	Mirador	AREAS - Engineering
EPR (Electronic Patient Record) - HL7 Integration possible?	yes	No	yes
EPR (Electronic Patient Record) - HL7 Version	ver 2.3		
EPR (Electronic Patient Record) - Non-HL7 Integration Patterns		Webservices, Facelink	
RIS-PACS - Name and Brand	CSC - Lorenzo Rad Centre/ Sectra	Sectra RIS/PACS	
RIS-PACS - HL7 Integration possible?	yes	Yes	
RIS-PACS - HL7 Version	ver 2.3 /2.31	2.4/2.5	
RIS-PACS - Non-HL7 Integration Patterns	DICOM	Webservices	
LIS - Name and Brand	Medat - David	Glims	
LIS - HL7 Integration possible?	yes	Yes	
LIS - HL7 Version	ver 2.2	2.5	
LIS - Non-HL7 Integration Patterns		Webservices	
Anatomic Pathology Laboratory System - Name and Brand	DC-Systeme - DC-Pathos	Sympathy	NOEMALIFE
Anatomic Pathology Laboratory System - HL7 Integration possible?	yes	Yes	no
Anatomic Pathology Laboratory System - HL7 Version	ver 2.2	2.5	n/a
Anatomic Pathology Laboratory System - Non-HL7 Integration Patterns		Webservices	PDF with meta data
Cross Document Repository - Name and Brand	No	Allgeier	

	HNO	VUMC	UNIPR
Hospitalization Ward (Involved in the ORAMOD Project)			
Cross Document Repository - HL7 Integration possible?		Yes	
Cross Document Repository - HL7 Version		3.0	
Cross Document Repository - Non-HL7 Integration Patterns		SQL Server	
Middleware - Name and Brand	Cloverleaf	BizTalk 2013	
Middleware - HL7 compatibility?	yes	Yes	
Tech Resources			
Do Hospitalization Ward and Outpatient Clinic have a wi-fi coverage? (pls specify the extension and bandwidth).	No	Yes	No
Do Hospitalization Ward and Outpatient Clinic have an Ethernet LAN coverage (pls specify the bandwidth)	Yes	Yes	Yes
Do you have a Test and Pre-production Environment?	Yes	Yes	Yes
Do you have a virtualization System? (e.g. VMWARE, KVM, etc.)	VMWARE	Yes	Yes
Do you have a Remote Access system and policy? (VPN?)	Yes	Yes	
Do you have a first level Help Desk? If yes, what is the daily and weekly service coverage.	Yes: 7-17 h, 5days/week	Yes	
Do you have a database management system? Which (Oracle, SQLServer, Postgress, ecc) and wich version?	Oracle, SQLServer	Yes. SQLServer, Oracle, FireBird, Progress, Sybase	
Number of Personal Computers or equivalent devices	4500	14000	
Amount of Technical Staff (units)	50	120	20
Do technical staff has J2EE platform and JAVA competence?	No	No	No
Do technical staff has DBMS competence? (pls specify the DBMS)	Oracle	Yes. SQLServer, Oracle, FireBird.	Yes. Oracle
Do technical staff has JBOSS platform and TOMCAT competence?	No	No.	Yes
Do you have and operate Linux OS servers? Which distribution and version?	Yes	Yes.	Yes

Table 2- Hospitals technical survey

2.3 The change that OraMod will bring to the hospital environment

In order to let different IT realities to be integrated one another, by an interoperability logical point of view, we present a series of figures that depict how the OraMod Platform will affect the single partners environments.

In each image the numbers connect to a standard, that is explained in the following chapter.

2.3.1 ORAMOD logical overview and standards keypoints

The first figure relates the main standard affecting each partner.

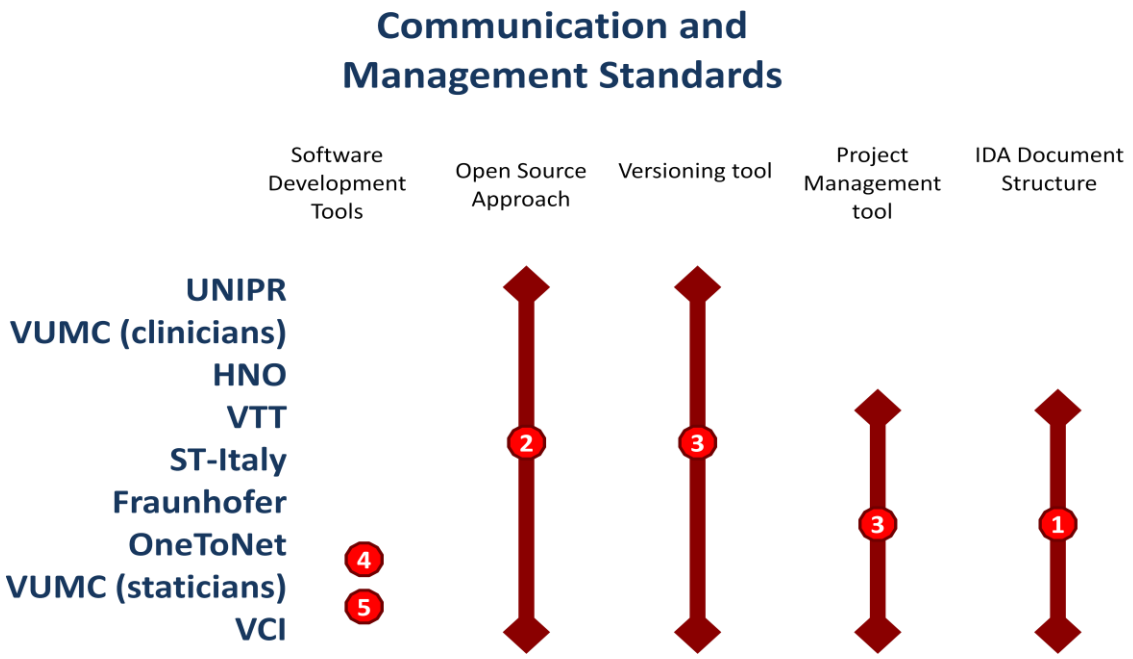


Fig. 1 - ORAMOD Project general Standards

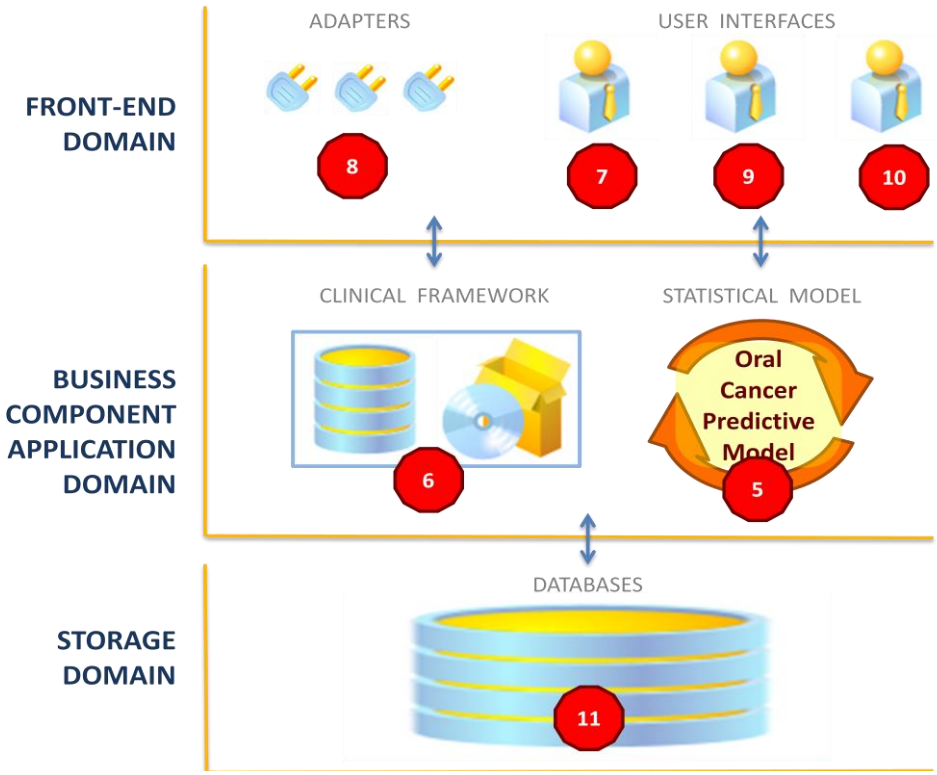


Fig. 2 - ORAMOD Architectural Standards

The previous standards refer to the different architectural layers of the project.

All the other standards are related to each interconnection (Fig. 3) and define how the various existing products will work with the core OraMod data bank.

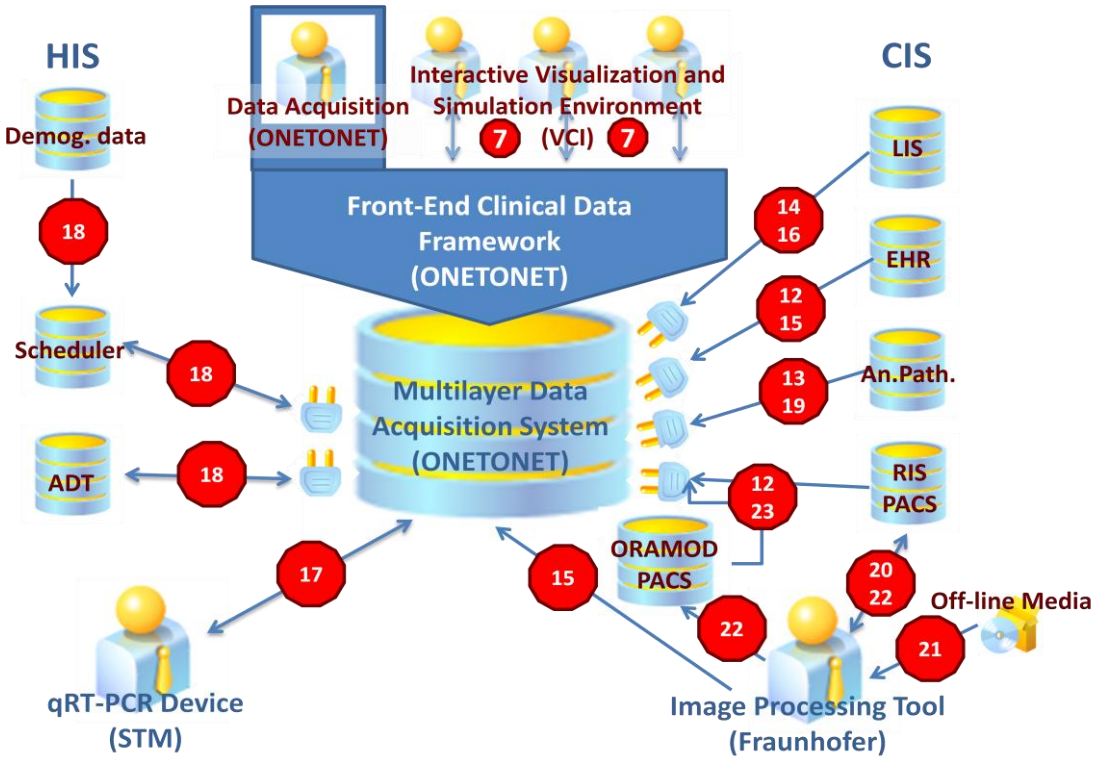


Fig. 3 - ORAMOD Interoperability Standards

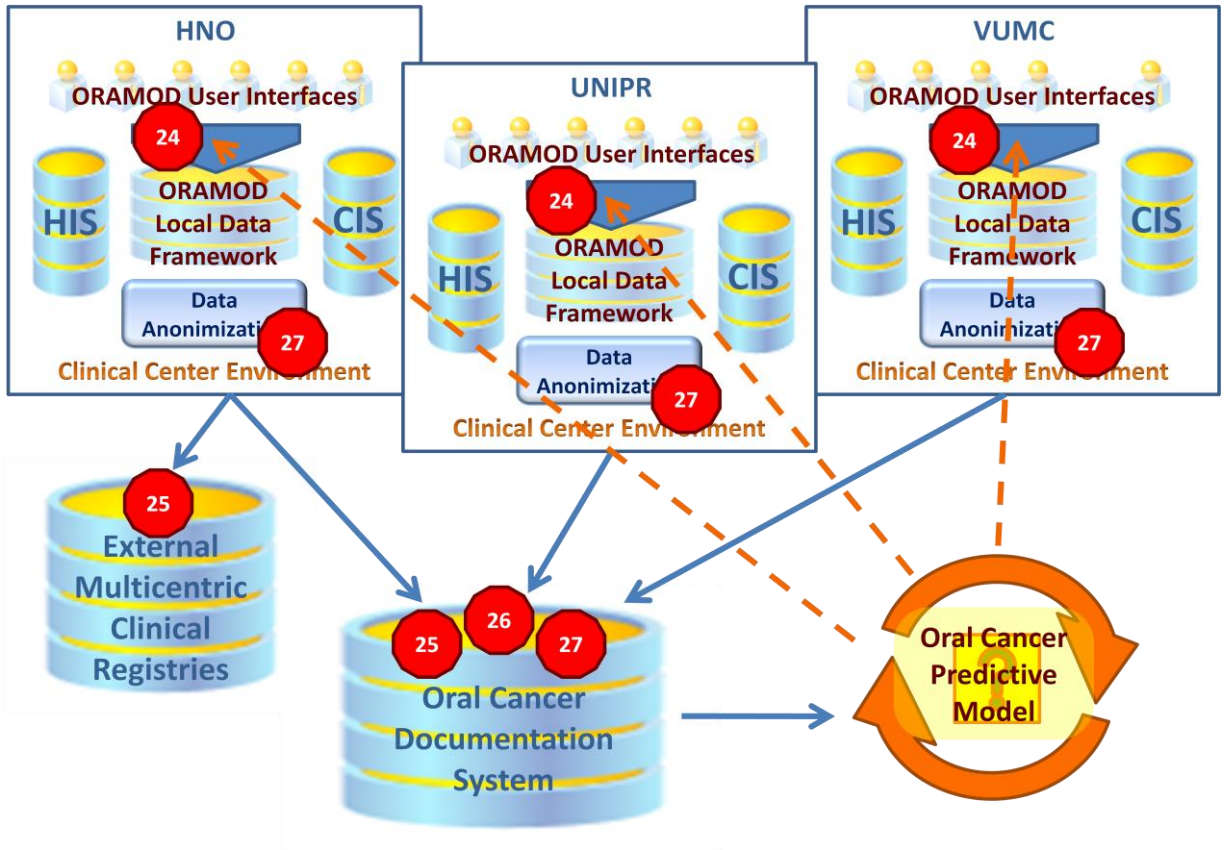


Fig. 4 - ORAMOD Overall Security Standards

The last image refers to the Medical Software Security Standards (cfr. Appendix VIII).

3 The technical standards identified for the project

#	Standard	use	why	Partner involved
1	IDA Programme [5]	It provides guidance and template material which is intended to assist the relevant management or technical staff, whether clinic or information technology personnel, in producing a project-specific Technical Design Document.	The IDA Programme identifies what can be referred to as a Reference Architecture for building blocks on which a variety of systems should be assembled. The Reference Architecture should then become part of the IDA Architecture Guidelines.	All the Partners producing technical analysis documents
2	Open Source Software Approach	**FOSS** will be adopted for Operating Systems, RDBMS, Developing and Management SW tools and General Productivity Suites	Be in accordance with EU general directives on Open Source Software usage	All the Partners
3	Project management Tools	**REDMINE** will be the open source web based Project management tool	REDMINE is an Open Source web based tool providing all the features requested for the project management of this project, such as: web accessibility, Project Wiki and Project internal forums, email notifications, time tracking, SVN integration, Multilanguage support	All the Partners
		SVN will be the ORAMOD Versioning tool used for both the Documents and the Shared Source Code	SVN (Subversion) is an Open Source universally recognized, centralized version control system fully compatible with the most common multiplatform versioning clients	All the technical Partners

#	Standard	use	why	Partner involved
4	Software Development Tools - Eclipse	ORAMOD original source code development, managing and deploying	Eclipse is the most common open source development platform and is already adopted by the code producing ORAMOD technical partners	ONETONET VCI
5	Software Development Tools - R Environment	Language and environment for statistical computing and graphics. Will be used to develop the predictive model	R is the open source language and environment adopted by VUMC statistician team	VUMC
6	JEE -EJB	Application layer developing language and specifications for the business component of the ORAMOD Project	EJB is the re-usable distributed architectural technology adopted by ONETONET framework and interoperability system EJB is also suggested by IDA programme as one of the technologies allowing the setting of "building blocks [...] forming the technical basis of the IDA developments	ONETONET
7	W3C JAVASCRIPT Web APIs	Interface language between the Clinical framework and the "virtual patient" user interfaces	JAVASCRIPT APIs allow developers to create a bridge between the browser and the platform it is running on, thus defining a more distinct but effectively dynamic and interactive borderline between 2 different applications or different tiers	ONETONET VCI

#	Standard	use	why	Partner involved
8	W3C - XML	XML will be the standard structured text-based format for all the structured data interoperability between different systems in the project. XML has been also adopted by more recent HL7 message formats.	XML is one of the most widely-used formats for sharing structured information today: between programs, between people, between computers and people, both locally and across networks	All the technical Partners
9	W3C Internationalization	All Web user interfaces will be compliant with W3C internationalization requirements		VCI
10	W3C Web Accessibility	All Web user interfaces will be compliant with W3C accessibility requirements		VCI
11	RDBMS	All the local databases. The oral cancer documentation system	PostgreSQL is the most powerful, open source enterprise class, object-relational database system, strongly conform to the ANSI-SQL:2008 standard and highly customizable.	ONETONET
12	IHE Cross-Enterprise Document Sharing Integration Profile	Interoperability scenarios between the ONETONET clinical framework and every single Clinical Partner HIS	World-Wide recognized Clinical Document Interoperability Scenario	ONETONET All the clinical Partners
13	IHE Anatomic Pathology Technical Framework	Interoperability scenarios between the ONETONET clinical framework and every single Clinical Partner HIS	World-Wide recognized Anatomic Pathology Interoperability Scenario	ONETONET All the clinical Partners
14	IHE Laboratory Technical Framework	Interoperability scenarios between the ONETONET clinical framework and every single Clinical Partner LIS	World-Wide recognized Laboratory Interoperability Scenario	ONETONET All the clinical Partners

#	Standard	use	why	Partner involved
15	HL7 CDA rel. 2	Interoperability data exchange messaging between the ONETONET clinical framework and every single Clinical Partner CIS	World-Wide recognized Clinical Data Interoperability Messaging Structure	ONETONET Fraunhofer All the clinical Partners
16	HL7 Clinical laboratory Automation	Interoperability data exchange messaging between the ONETONET clinical framework and every single Clinical Partner LIS	World-Wide recognized Laboratory Data Interoperability Messaging Structure	ONETONET All the clinical Partners
17	HL7 Genomic Data Exchange	Interoperability data exchange messaging between the ONETONET clinical framework and ST-Italy qt-PCR device	World-Wide recognized Genomic Data Interoperability Messaging Structure	ONETONET ST-Italy
18	HL7 RIM – ISOHL7 21731	Interoperability data exchange messaging between the ONETONET clinical framework and every single Clinical Partner HIS and CIS	World-Wide recognized Health and Administrative Data Interoperability Messaging Structure	ONETONET All the clinical Partners
19	LOINC	Interoperability between the ONETONET clinical framework and Clinical Partner LIS, Anatomic Pathology and Genomic data systems	World-Wide recognized universal standard for identifying medical laboratory observations	ONETONET All the clinical Partners
20	DICOM Query/Retrieve	To find lists of images based on some criteria's and to select them	DICOM is the World-Wide recognized universal standard for Imaging	Fraunhofer All the clinical Partners
21	DICOM Off-line Media	It describes how to store medical imaging information on removable media	DICOM is the World-Wide recognized universal standard for Imaging	Fraunhofer
22	DICOM Store &Storage commitment	Used to send images to a PACS or workstation	DICOM is the World-Wide recognized universal standard for Imaging	Fraunhofer ONETONET

#	Standard	use	why	Partner involved
23	DICOM Modality Worklist	Service to obtain patient details and worklist of his exams	DICOM is the World-Wide recognized universal standard for Imaging	Fraunhofer ONETONET All the clinical Partners
24	Mobile Medical Applications		EU has not a specific directive on the Health IT, so we will take care the following USA FDA main directives	All the technical Partners
25	Connected Health			All the technical Partners
26	Cybersecurity			All the technical Partners
27	Confidentiality of Data			All the technical Partners

Table 3 - OraMod technical standards

Appendix I - IDA Programme Document Specifications

Overview

The first phase of the programme (IDA I), which started in 1995 (Decision 95/468/EC), contributed to the establishment of large telematic networks in the areas of employment, health, agriculture, statistics and competition.

Areas of intervention

The IDA II programme supports the implementation of projects of common interest relating in particular to the development and creation of telematic networks in the area of Community policies such as Economic and Monetary Union (EMU), consumer protection, health and transport.

Objectives

The objectives pursued by the Community with the IDA programme are:

- to achieve a high degree of interoperability between the telematic networks in the Member States and between the Community and the Member States;
- to make such networks converge towards a common telematic interface between the Community and the Member States;
- to achieve benefits for Member State administrations and the Community resulting in particular from the streamlining of operations, a reduction in maintenance, speeding up the implementation of new networks and the provision of safe and reliable data interchange;
- to extend the benefits of these networks to EU businesses and citizens;
- to promote the spread of best practice and encourage the development of innovative telematic solutions in administrations.

Eligibility criteria

Projects must be in the area of Community policies and activities.

Priority is given to projects improving the economic viability of public administrations, European institutions, the Member States and regions which, by setting up or developing a sectoral network:

- help to overcome the obstacles to the free movement of goods, persons, services and capital;
- contribute to the successful implementation of EMU;
- promote institutional cooperation between the Community institutions and between those institutions and the national and regional administrations;
- help to safeguard the financial interests of the Community and Member States and combat fraud;
- contribute to preparing for EU enlargement;

- promote the competitiveness of Community industry and, more particularly, small and medium-sized businesses;
- benefit EU citizens.

Beneficiaries

The principal beneficiaries are the national or regional administrations of the Member States and the Community institutions.

Link to the eEurope Action Plan

In January 2002 the IDA programme also became the instrument for implementing the e-government chapter of the eEurope 2005 action plan. The IDA programme supports the development of services geared to safe and efficient electronic data interchange between the various levels of administration. These are essential for providing modern on-line public services, as provided for in the eEurope 2005 Action Plan.

Under the IDA programme, the Commission has set up and funded for the last two years a secure network communications infrastructure for data interchange (TESTA - Trans-European Services for Telematics between Administrations) between practically all the administrations of the Member States - and soon of the accession countries as well - and the European institutions. With the development of e-government projects, the TESTA network will be able to support pan-European services for citizens and businesses. IDA also funds the TESS (Telematics in Social Security) programme.

Basis of this document

This following introductory sections set out an approach to designing systems that may be developed under IDA. It attempts to set standards and create a consistent approach to the design and development of systems across the IDA Programme. It will enable the Programme to benefit from 'economies of scale' and a consistency in the approach to building and deploying systems. Important issues that need to be considered include the architecture of systems, links to legacy systems, contemporary approaches to design (Object Oriented Design), aims for code re-use and the need to develop systems that will work on an operational basis over many years and the associated desire to make such systems easily supportable and affordable.

A key point will be to build on the work already carried out in IDA and its predecessor programmes, where a large number of specific 'technical' developments were undertaken looking at, for example, standards for data exchange, such as GESMES, and the introduction of contemporary technologies and infrastructures.

The concept of a **Reference Architecture** is also introduced as part of the process of creating an interoperable environment which facilitates the exchange of information between administrations through setting out a number of standard building blocks around which solutions can be assembled. These building blocks or 'components reflect the emerging technologies that should form the technical basis of the IDA

developments. These include the Common Object Request Broker Architecture (CORBA), Remote Method Invocation (RMI) and Enterprise JavaBeans (EJB) technologies that highlight code re-use, scalability and the creation of interoperable architectures around legacy environments.

A Reference Architecture for the IDA Programme

One of the greatest difficulties facing the IDA Programme is how to create an interoperable architecture across European administrations whilst allowing individual projects to manage their own IDA developments with sub-delegated authority. Under sub-delegation, it would be possible for sectors seeking to be selected to build elements of the overall IDA Programme to propose their own solutions, or the solutions proposed by their selected suppliers, thereby exacerbating the existing heterogeneous environment, which is in direct conflict with the aim of creating an interoperable architecture.

In practice the IDA Programme should seek to create what can be referred to as a Reference Architecture for systems – in effect a series of established building blocks on which a variety of systems should be assembled. The Reference Architecture should then become part of the IDA Architecture Guidelines[2].

The Reference Architecture would take an n-tier, client-server model as its basis, and state that the persistence layer would be implemented through a number of standard API calls to a particular database system. This use of a Reference Architecture would provide some guidelines to organisations tasked with developing individual elements of the IDA Programme and would, in time, move solutions towards a common homogeneous standard – with all of its associated benefits.

Recommendations could be made on the reporting tool that should be used in systems if one were required, the Web Browser could be standardized, as could any security features that needed to be built into the system. The Reference Architecture would facilitate re-use throughout the IDA Programme as individual projects could be arranged to support the development of libraries of components that can be re-used. This will be one of the major potential benefits that could arise from the creation of a Reference Architecture.

[2] The Architecture Guidelines were developed in 1999 in co-operation with the TAG (Telematics for Administrations Group) Subgroup on Horizontal and Legal Issues, a group of experts from EU Member States and EEA states. They describe concepts and references for a well-defined, common architecture that supports network interoperability within and across different administrations. Article 4 of the IDA "Interoperability" Decision (No. 1720/1999/EC) defines the requirement to develop the Guidelines as one of IDA the IDA Horizontal Actions and Measures. Article 5 of the IDA "Guidelines" Decision (No. 1719/1999/EC) requires projects to make use of them. An update to the Guidelines is expected in early 2001.

Appendix II - Open Source Software Approach

OneToNet, while identifying the most appropriate tools for the ORAMOD project management and technical collaboration, had 2 cornerstones: be fully in accordance with the E.U. general directives of open source software usage (Ref. Ares (2011)1016907 - 26/09/2011) [1], and get full advantage of the multichannel capabilities of the Web.

Inter-Institutional Committee for Informatics- Conclusions of the discussion on Open Source Software and on the alignment of Open Source Software strategies

The mandate of the Inter-Institutional Committee for Informatics (hereafter "CII") includes, on the one hand, the exchange of information about the IT policies of the Institutions represented in it (hereafter "the Institutions") and, on the other hand, the identification and encouragement of potential areas of synergy.

In this context, during its meeting held in Brussels on 14 December 2010, the CII held a discussion about the usage of Open Source Software (OSS) in the Institutions, based on a survey distributed ahead of the meeting and filled in by all the Institutions as well as by 15 additional EU Agencies (hereafter "the survey").

In addition, in order to initiate and facilitate the discussion, the Commission presented the latest update to its Strategy for the internal use of OSS, which it has maintained since 2001, a summary of which is published on the Internet.

The findings of the survey were the following:

1. *OSS is, in general, already included in public procurement processes in the Institutions, with OSS solutions being considered at the same level as proprietary ones as part of the product management procedures in place. Decisions are taken on the basis of fitness for purpose and Total Cost of Ownership (TCO), including protection of investments.*
2. *As a result, OSS products are already represented in the software portfolio of most Institutions. For example:*
    a. A majority of Institutions already offer formal hosting services based on OSS stacks in their Data Centers in domains such as networking and operating systems.
    b. A majority of Institutions already use a large number of OSS solutions in the middleware layer, such as application servers and databases.
    c. In the area of information system development, OSS is widely used by development teams (typically in the form of OSS frameworks, libraries and add-on tools).
3. *By contrast, OSS products are not yet currently used widely in the Institutions in other areas, in particular:*
    a. In the collaboration and content management domain, where the OSS market is promising but very fragmented, although promising inroads have been made in areas such as the direct communication with citizens;
    b. In the end-user desktop environment, where (with the exception of the web browsers) OSS solutions do not yet offer, as of today, all the functionality required to make them viable alternatives for organizations of the complexity of the Institutions and/or would require significant migration investments.
4. *The patterns of use of OSS products in the Institutions are in line with those observed in the market, and in particular for large public administrations.*

5. *Some Institutions have acquired considerable experience with other aspects of the OSS approach, such as OSS-licensing (EUPL), community-building and promoting a culture of collaboration. These are not trivial tasks, as they require investment and careful planning, and may also have legal implications.*

6. *Until now, however, only a minority of the Institutions have formalized their strategy for the internal use of OSS.*

Following the discussion held during the meeting, the Institutions agreed:

1. *To reaffirm their commitment to endeavor that their IT infrastructures support all relevant standards1 adopted by standardization bodies as well as, to the widest possible extent, other widely used standards, regardless of the type of software used internally.*

2. *To declare that they will continue to consider OSS solutions in their public procurement processes on an equal footing with proprietary software, based on value for money (considering in particular the Total Cost of Ownership), and in full respect of the applicable procurement legislation.*

3. *To declare that, in order to align their practices in this area, they will adopt, where necessary, formal documents describing their strategies for the internal use of OSS, similar to the one presented by the European Commission during the meeting of 14 December 2010, and to those already adopted by some other Institutions. For the sake of transparency, these documents will be made public gradually, in full or in abridged form, on the Institutions' websites.*

http://osepa.eu/

Open Source software usage by European Public Administrations - Policy Recommendation Paper, 3rd version [2]**.**

OSEPA - Policy Recommendation Paper

Based on policy review and analysis, OSEPA highlights 25 recommendations on policy initiatives and actions, grouped in five broad FOSS policy areas as defined in the previous section:

4. *Data openness and reusability*

5. *Licensing, procurement and software market policies*

6. *FOSS adoption, integration and sustainability*

7. *Research & innovation*

8. *Training and education*

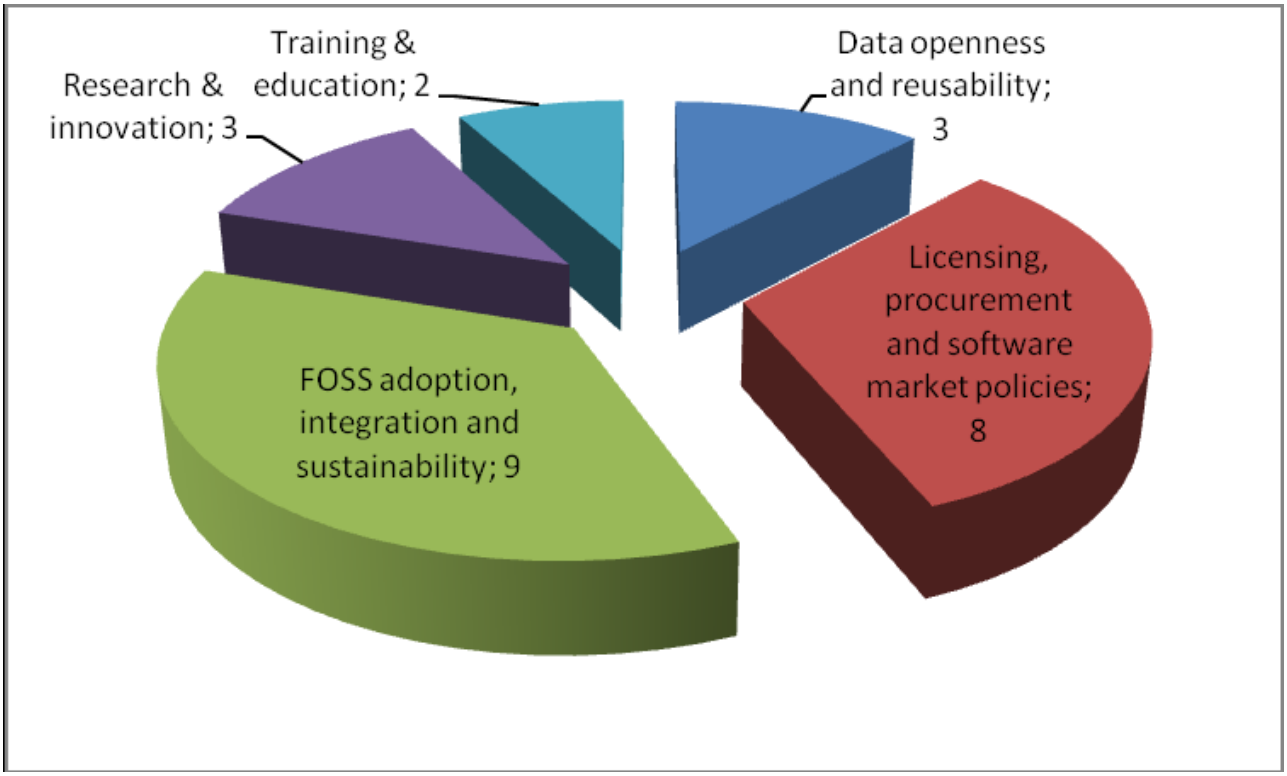


Fig. 5 - Proposed actions by policy area (Total=25)

Open Source software usage by European Public Administrations - Good Practice Guide covering various aspects of FOSS usage by European Public Administrations. 2nd (final) version.

OSEPA - CP3 Good Practice Guide. 2nd version.

FOSS European and National Policies and practices: Analysis and Recommendations.

OSEPA – Report of evidence of national FOSS policies. v6.

Appendix III - Software Tools

Project Management Tools

OneToNet identified **Redmine** (http://www.redmine.org/) [6] as the ORAMOD project management web application. The ORAMOD Project Management features and scenarios that Redmine is fully satisfying are referred in the following list:

- Flexible role based access control
- Gantt chart and calendar
- News, documents & files management
- Feeds & email notifications
- Project wiki
- Project forums
- Time tracking
- SCM integration (SVN, CVS, Git, Mercurial, Bazaar and Darcs)
- Issue creation via email
- Multilanguage support

In the advanced steps of the project, Redmine will also be used as the issue tracking system in order to provide to the clinicians a direct and fast communication tool with the developers. Redmine is an open source project and is released under the terms of the GNU General Public License v2 (GPL).

For documents and code versioning purposes, OneToNet identified **Apache™ Subversion®** (http://subversion.apache.org/) [7]: the Apache Software Foundation open source version control system.

Subversion is universally recognized and adopted as an open-source, centralized version control system characterized by its reliability, the simplicity of its model and usage, and its ability to support the needs of a wide variety of users and projects, from individuals to large-scale enterprise operations.

Subversion can be used with a large number of opensource, multiplatform versioning clients such as TortoiseSVN (http://tortoisesvn.net/), RabbitVCS (http://rabbitvcs.org/), Subcommander (http://subcommander.tigris.org/) and many others.

Software Development Tools

Eclipse

Eclipse is the certified tool [8].

Eclipse is a community for individuals and organizations who wish to collaborate on commercially-friendly open source software. Its projects are focused on building an open development platform comprised of extensible frameworks, tools and runtimes for building, deploying and managing software across the lifecycle. The Eclipse Foundation is a not-for-profit, member supported corporation that hosts the **Eclipse projects** and helps cultivate both an open source community and an ecosystem of complementary products and services.

The Eclipse Project was originally created by IBM in November 2001 and supported by a consortium of software vendors. The Eclipse Foundation was created in January 2004 as an independent not-for-profit

corporation to act as the steward of the Eclipse community. The independent not-for-profit corporation was created to allow a vendor neutral and open, transparent community to be established around Eclipse. Today, the Eclipse community consists of individuals and organizations from a cross section of the software industry.

*The Eclipse Foundation is funded by annual dues from our **members** and governed by a **Board of Directors**. Strategic Developers and Strategic Consumers hold seats on this Board, as do representatives elected by Add-in Providers and Open Source committers. The Foundation employs a full-time **professional staff** to provide services to the community but does not employ the open source developers, called committers, which actually work on the Eclipse projects. Eclipse committers are typically employed by organizations or are independent developers that volunteer their time to work on an open source project.*

In general, the Eclipse Foundation provides four services to the Eclipse community:

*1) **IT Infrastructure**,*

*2) **IP Management,***

*3) **Development Process**, and*

*4) **Ecosystem Development**.*

Full-time staff are associated with each of these areas and work with the greater Eclipse community to assist in meeting the needs of the stakeholders.

R

R [9] is a language and environment for statistical computing and graphics. It is a GNU project which is similar to the S language and environment which was developed at Bell Laboratories (formerly AT&T, now Lucent Technologies) by John Chambers and colleagues. R can be considered as a different implementation of S. There are some important differences, but much code written for S runs unaltered under R.

R provides a wide variety of statistical (linear and nonlinear modelling, classical statistical tests, time-series analysis, classification, clustering, ...) and graphical techniques, and is highly extensible. The S language is often the vehicle of choice for research in statistical methodology, and R provides an Open Source route to participation in that activity.

One of R's strengths is the ease with which well-designed publication-quality plots can be produced, including mathematical symbols and formulae where needed. Great care has been taken over the defaults for the minor design choices in graphics, but the user retains full control.

R is available as Free Software under the terms of the Free Software Foundation's GNU General Public License in source code form. It compiles and runs on a wide variety of UNIX platforms and similar systems (including FreeBSD and Linux), Windows and MacOS.

The R environment

R is an integrated suite of software facilities for data manipulation, calculation and graphical display. It includes

- an effective data handling and storage facility,
- a suite of operators for calculations on arrays, in particular matrices,
- a large, coherent, integrated collection of intermediate tools for data analysis,
- graphical facilities for data analysis and display either on-screen or on hardcopy, and
- a well-developed, simple and effective programming language which includes conditionals, loops, user-defined recursive functions and input and output facilities.

The term "environment" is intended to characterize it as a fully planned and coherent system, rather than an incremental accretion of very specific and inflexible tools, as is frequently the case with other data analysis software.

R, like S, is designed around a true computer language, and it allows users to add additional functionality by defining new functions. Much of the system is itself written in the R dialect of S, which makes it easy for users to follow the algorithmic choices made. For computationally-intensive tasks, C, C++ and Fortran code can be linked and called at run time. Advanced users can write C code to manipulate R objects directly.

Many users think of R as a statistics system. We prefer to think of it of an environment within which statistical techniques are implemented. R can be extended (easily) via *packages*. There are about eight packages supplied with the R distribution and many more are available through the CRAN family of Internet sites covering a very wide range of modern statistics.

R has its own LaTeX-like documentation format, which is used to supply comprehensive documentation, both on-line in a number of formats and in hardcopy.

Appendix IV - Web-based Distributed Architecture

To fulfill the project requirements, and mainly to achieve the two main goals of this segment of the project (i.e. a distributed and interoperable platform in the clinical environment) we will particularly focus on architecture standards and on integration standards. In particular we will analyze the standards related to web-based distributed architecture based on EJB (Enterprise Java Beans), such as Oracle J2EE (Java 2 Enterprise Edition) standards [10], as well as all the major W3C standards and rules about web browser interoperability, accessibility (W3C Web Accessibility Initiative – WAI), internationalization (Internationalization -I18n Activity).

JAVA EE

Enterprise Java Beans

Enterprise JavaBeans (EJB) technology is the server-side component architecture for Java Platform, Enterprise Edition (Java EE). EJB technology enables rapid and simplified development of distributed, transactional, secure and portable applications based on Java technology.

EJB 3.0 Specification Final Release This specification defines the new simplified EJB API targeted at ease of development. It also includes the new Java Persistence API for the management of persistence and object/relational mapping with Java EE and Java SE.

Java Persistence API The Java Persistence API is the standard API for the management of persistence and object/relational mapping. It provides an object/relational mapping facility for application developers using a Java domain model to manage a relational database. The Java Persistence API is part of the Java EE platform. It can also be used in Java SE environments.

EJB 2.1 specification This spec, created under the Java Community Process (JCP), enhances EJB architecture with support for Web services, making it easier to implement and deploy Web services applications based on Java technology.

W3C Standards

JAVASCRIPT Web APIs [11]

While the most common scripting language ECMA script (more widely known as JavaScript) is developed by Ecma, a great many of the APIs made available in browsers have been defined at W3C.

What is scripting?

A script is program code that doesn't need pre-processing (e.g. compiling) before being run. In the context of a Web browser, scripting usually refers to program code written in JavaScript that is executed by the browser when a page is downloaded, or in response to an event triggered by the user.

Scripting can make Web pages more dynamic. For example, without reloading a new version of a page it may allow modifications to the content of that page, or allow content to be added to or sent from that page. The former has been called DHTML (Dynamic HTML), and the latter AJAX (Asynchronous JavaScript and XML).

Beyond this, scripts increasingly allow developers to create a bridge between the browser and the platform it is running on, making it possible, for example, to create Web pages that incorporate information from the user's environment, such as current location, address book details, etc.

This additional interactivity makes Web pages behave like a traditional software application. These Web pages are often called *Web applications* and can be made available either directly in the browser as a Web page, or can be packaged and distributed as *Widgets*.

What scripting interfaces are available ?

The most basic scripting interface developed at W3C is the DOM, the Document Object Model which allows programs and scripts to dynamically access and update the content, structure and style of documents. DOM specifications form the core of DHTML.

Modifications of the content using the DOM by the user and by scripts trigger events that developers can make use of to build rich user interfaces.

A number of more advanced interfaces are being standardized, for instance:

- XMLHttpRequest makes it possible to load additional content from the Web without loading a new document, a core component of AJAX,
- the Geolocation API makes the user's current location available to browser-based applications,
- several APIs make the integration of Web applications with the local file system and storage seamless.

WAI ARIA offers mechanisms to ensure that this additional interactivity remains usable independent of devices and disabilities. Additional considerations apply to the development of Web applications for mobile devices.

Beyond scripting

While scripting offers a great opportunity to develop new interfaces and experiment with new user interactions, over time a number of these additions benefit from a more *declarative* approach; for instance, instead of having each and every developer re-implement a calendar-interface that allows a user to pick a date, defining an input type (`<input type='date' />`) that does it automatically saves a lot of time and bugs, and creates a ground for further innovation.

Beyond the set of declarative interfaces made available through HTML, several technologies have been developed to make these Declarative Web Applications possible.

Access to the Web for all has been a fundamental concern and goal of the World Wide Web Consortium since the beginning. Unfortunately, it is easy to overlook the needs of people from cultures different to your own, or who use different languages or writing systems. If you do, you will build specifications and content that present barriers to the use of your technology or content for many people around the world.

XML [12]

What is XML?

The Extensible Markup Language (XML) is a simple text-based format for representing structured information: documents, data, configuration, books, transactions, invoices, and much more. It was derived from an older standard format called SGML (ISO 8879), in order to be more suitable for Web use.

What is XML Used For?

XML is one of the most widely-used formats for sharing structured information today: between programs, between people, between computers and people, both locally and across networks.

A short example:

```
<part number="1976">
  <name>Windscreen Wiper</name>
  <description>The Windscreen wiper
    automatically removes rain
    from your windscreen, if it
    should happen to splash there.
    It has a rubber <ref part="1977">blade</ref>
    which can be ordered separately
    if you need to replace it.
  </description>
</part>
```

If you are already familiar with HTML, you can see that XML is very similar. However, the syntax rules of XML are strict: XML tools will not process files that contain errors, but instead will give you error messages so that you fix them. This means that almost all XML documents can be processed reliably by computer software.

The main differences from HTML are:

1. All elements must be *closed* or marked as *empty*.
2. Empty elements can be closed as normal, `<happiness></happiness>` or you can use a special short-form, `<happiness />` instead.
3. In HTML, you only need to quote an attribute value under certain circumstances (it contains a space, or a character not allowed in a name), but the rules are hard to remember. In XML, attribute values must always be quoted:
   `<happiness type="joy" />`
4. In HTML there is a built-in set of element names (along with their attributes). In XML, there are no built-in names (although names starting with `xml` have special meanings).
5. In HTML, there is a list of some built-in character names like `é` for é but XML does not have this. In XML, there are only five built-in character entities: <, >, &, " and ' for <, >, &, " and ' respectively. You can define your own entities in a Document Type Definition, or you can use any Unicode character (see next item).
6. In HTML, there are also numeric character references, such as & for &. You can refer to any Unicode character, but the number is decimal, whereas in the Unicode tables the number is usually in hexadecimal. XML also allows hexadecimal references: & for example.

XML has a number of advantages over many other formats. For any particular scenario, you might be able to come up with a better format, but then you would have to include costs of converting and processing

your format, and of training, and of the XML-specific editing and searching tool that are now very widely available. Some of the advantages of XML include:

Redundancy

> XML markup is very verbose. For example, every end tag must be supplied, such as `</description>` in the example. This lets the computer catch common errors such as incorrect nesting.

Self-describing

> The readability of XML (it is a text-based format) and the presence of element and attribute names in XML means that people looking at an XML document can often get a head start on understanding the format (and it also helps people to find mistakes!)

Network effect and the XML Promise

> Any XML document can be read and processed by any XML tool whatsoever. Of course, some XML tools might want specific XML markup, but the XML format itself can be read by any XML parser: you can't say, this XML document is only to be processed by such-and-such a tool.
>
> This means that every new XML document increases the value of every other XML document, and of every XML tool, and every new XML tool increases the value of every XML document and hence of every other tool. Today, XML is the most widely-used format of its kind anywhere in the world.

Internationalization [13]

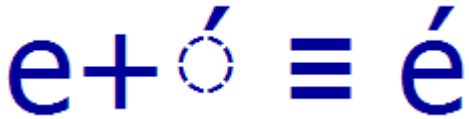
What is Internationalization?

If you internationalize, you design or develop your content, application, specification, and so on, in a way that ensures it will work well for, or can be easily adapted for, users from any culture, region, or language.

The word 'Internationalization' is often abbreviated to 'i18n'. This is widely used abbreviation, derived from the fact that there are 18 letters between the 'i' and the 'n'.

Examples

One fundamental aspect of internationalization is to ensure that the technology supports text in any writing system of the world. This is why W3C technologies are built on the universal character set, Unicode. It may be necessary to also support other legacy character sets and encodings.



There are other factors to consider, however, when using characters. For example, Unicode based encodings allow the exact same text to be stored using slightly different combinations of characters. For

efficiency and accuracy in comparing, sorting and parsing text, the different sequences need to be recognised as 'canonically equivalent'. You need to consider how to manage this when developing applications or specifications that perform or rely on such tasks.

国际化活动、万维网联盟

Sometimes different writing systems require special support. For example, Japanese, Chinese, Korean and Mongolian can be written vertically, so the W3C is ensuring CSS, SVG and XSL-FO will allow for vertical text support. Text alignment and justification methods are also different for such scripts, and different again for scripts like Thai and Tibetan. Other local typographic conventions often exist for such things as emphasis, annotations, list numbering, and the like. These typographic approaches need to be supported in style sheets.

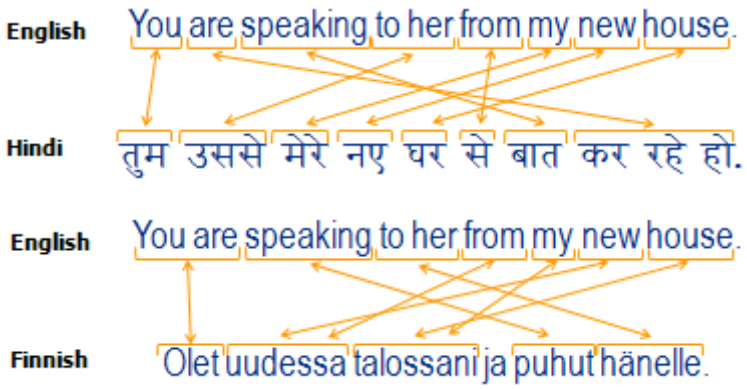
Arabic, Hebrew, Persian, Urdu and similar languages mix right-to-left and left-to-right text on the same line, and it is important to be able to control the direction of the surrounding context for that to work properly. This means that schema and format developers need to provide ways for authors to control direction in their content. Schemas, markup languages and formats should also support a number of other constructs needed for efficient handling of content during translation and localisation.

The title is مفتاح معايير الويب in Arabic.

If you are dealing with HTML forms or designing ontologies relating to people's names and addresses, you will need to consider how to enable the many different approaches to formatting data that are possible around the world. You may also need to support alternative calendars, time zones and daylight savings, names and addresses in both native plus transliterated forms, etc.

Content developers and content management systems must also be prepared to deal with linguistic and cultural issues. For example, a sentence that is constructed by combining several phrases together in one language may be impossible to translate sensibly in a language with a different sentence structure. For

example, in the Japanese translation of "Page 1 of 34" all elements in the phrase would be in reverse order. Your application must not restrict the order in which these elements can be combined. Specifications for technologies such as widgets and voice browsers should also avoid locking developers into an English-biased syntax for such things as composing messages or firing events associated with text.



Cultural problems also need to be considered. Symbolism can be culture-specific. The check mark means correct or OK in many countries. In some countries, however, such as Japan, it can be used to mean that something is *incorrect*. Japanese localizers may need to convert check marks to circles (their symbol for 'correct') as part of the localization process.

These are just a few examples of many. The key message is that design (whether it be of a markup language, a protocol, a content management system, a widget or application, etc.) needs to be flexible enough to accommodate local needs.

Web Accessibility [14]

What is Web Accessibility

Web accessibility means that people with disabilities can use the Web. More specifically, Web accessibility means that people with disabilities can perceive, understand, navigate, and interact with the Web, and that they can contribute to the Web. Web accessibility also benefits others, including older people with changing abilities due to aging.

Web accessibility encompasses all disabilities that affect access to the Web, including visual, auditory, physical, speech, cognitive, and neurological disabilities. The document "**How People with Disabilities Use the Web**" **Error! Reference source not found.** describes how different disabilities affect Web use and includes scenarios of people with disabilities using the Web.

Millions of people have disabilities that affect their use of the Web. Currently most Web sites and Web software have **accessibility barriers** that make it difficult or impossible for many people

with disabilities to use the Web. As more accessible Web sites and software become available, people with disabilities are able to use and contribute to the Web more effectively.

Web accessibility also **benefits** people *without* disabilities. For example, a key principle of Web accessibility is designing Web sites and software that are flexible to meet different user needs, preferences, and situations. This **flexibility** also benefits people *without* disabilities in certain situations, such as people using a slow Internet connection, people with "temporary disabilities" such as a broken arm, and people with changing abilities due to aging. The document "Developing a Web Accessibility Business Case for Your Organization" **Error! Reference source not found.** describes many different benefits of Web accessibility, including **benefits for organizations**.

Making the Web Accessible

Much of the focus on Web accessibility has been on the responsibilities of Web developers. However, Web software also has a vital role in Web accessibility. Software needs to help developers produce and evaluate accessible Web sites, and be usable by people with disabilities.

One of the roles of the Web Accessibility Initiative (WAI) is to develop guidelines and techniques that describe accessibility solutions for Web software and Web developers. These WAI guidelines are considered the international standard for Web accessibility.

The document "**Essential Components of Web Accessibility**" describes the different Web accessibility roles, and how specific improvements could substantially advance Web accessibility.

Making Your Web Site Accessible

Making a Web site accessible can be simple or complex, depending on many factors such as the type of content, the size and complexity of the site, and the development tools and environment.

Many accessibility features are easily implemented if they are planned from the beginning of Web site development or redesign. Fixing inaccessible Web sites can require significant effort, especially sites that were not originally "coded" properly with standard XHTML markup, and sites with certain types of content such as multimedia.

The document "Implementation Plan for Web Accessibility" lists basic steps for addressing accessibility in Web projects. The Web Content Accessibility Guidelines and techniques documents provide detailed information for developers.

Evaluating the Accessibility of a Web Site

When developing or redesigning a site, evaluating accessibility early and throughout the development process can identify accessibility problems early when it is easier to address them. Simple techniques such as changing settings in a browser can determine if a Web page meets some accessibility guidelines. A comprehensive evaluation to determine if a site meets all accessibility guidelines is much more complex.

There are evaluation tools that help with evaluation. However, no tool alone can determine if a site meets accessibility guidelines. Knowledgeable **human evaluation is required** to determine if a site is accessible.

The document "Evaluating Web Sites for Accessibility" provides guidance on preliminary reviews using techniques to quickly assess some of the accessibility problems on a site. It also provides general procedures and tips for evaluating conformance to accessibility guidelines.

For More Information

The WAI Web site provides guidelines and resources to help make the Web accessible. These range from very short summaries, such as "Quick Tips to Make Accessible Web Sites," to resources on managing accessibility, to detailed technical references.

Related resources for making the Web accessible are also available from other organizations, and many can be found on the Web.

Appendix V - RDBMS

PostgreSQL

PostgreSQL is a powerful, open source object-relational database system. It has more than 15 years of active development and a proven architecture that has earned it a strong reputation for reliability, data integrity, and correctness. It runs on all major operating systems, including Linux, UNIX (AIX, BSD, HP-UX, SGI IRIX, Mac OS X, Solaris, Tru64), and Windows. It is fully ACID compliant, has full support for foreign keys, joins, views, triggers, and stored procedures (in multiple languages). It includes most SQL:2008 data types, including INTEGER, NUMERIC, BOOLEAN, CHAR, VARCHAR, DATE, INTERVAL, and TIMESTAMP. It also supports storage of binary large objects, including pictures, sounds, or video. It has native programming interfaces for C/C++, Java, .Net, Perl, Python, Ruby, Tcl, ODBC, among others, and exceptional documentation.

An enterprise class database, PostgreSQL boasts sophisticated features such as Multi-Version Concurrency Control (MVCC), point in time recovery, tablespaces, asynchronous replication, nested transactions (savepoints), online/hot backups, a sophisticated query planner/optimizer, and write ahead logging for fault tolerance. It supports international character sets, multibyte character encodings, Unicode, and it is locale-aware for sorting, case-sensitivity, and formatting. It is highly scalable both in the sheer quantity of data it can manage and in the number of concurrent users it can accommodate. There are active PostgreSQL systems in production environments that manage in excess of 4 terabytes of data. Some general PostgreSQL limits are included in the table below.

Limit	Value
Maximum Database Size	Unlimited
Maximum Table Size	32 TB
Maximum Row Size	1.6 TB
Maximum Field Size	1 GB
Maximum Rows per Table	Unlimited
Maximum Columns per Table	250 - 1600 depending on column types
Maximum Indexes per Table	Unlimited

Table 4 - PostgreSQL limits

PostgreSQL has won praise from its users and industry recognition, including the Linux New Media Award for Best Database System and five time winner of the The Linux Journal Editors' Choice Award for best DBMS.

Featureful and Standards Compliant

PostgreSQL prides itself in standards compliance. Its SQL implementation strongly conforms to the ANSI-SQL:2008 standard. It has full support for subqueries (including subselects in the FROM clause), read-committed and serializable transaction isolation levels. And while PostgreSQL has a fully relational system

catalog which itself supports multiple schemas per database, its catalog is also accessible through the Information Schema as defined in the SQL standard.

Data integrity features include (compound) primary keys, foreign keys with restricting and cascading updates/deletes, check constraints, unique constraints, and not null constraints.

It also has a host of extensions and advanced features. Among the conveniences are auto-increment columns through sequences, and `LIMIT/OFFSET` allowing the return of partial result sets. PostgreSQL supports compound, unique, partial, and functional indexes which can use any of its B-tree, R-tree, hash, or GiST storage methods.

GiST (*Generalized Search Tree*) indexing is an advanced system which brings together a wide array of different sorting and searching algorithms including B-tree, B+-tree, R-tree, partial sum trees, ranked B+-trees and many others. It also provides an interface which allows both the creation of custom data types as well as extensible query methods with which to search them. Thus, GiST offers the flexibility to specify *what* you store, *how* you store it, and *the ability to define new ways* to search through it --- ways that far exceed those offered by standard B-tree, R-tree and other generalized search algorithms.

GiST serves as a foundation for many public projects that use PostgreSQL such as OpenFTS and PostGIS. OpenFTS (Open Source Full Text Search engine) provides online indexing of data and relevance ranking for database searching. PostGIS is a project which adds support for geographic objects in PostgreSQL, allowing it to be used as a spatial database for geographic information systems (GIS), much like ESRI's SDE or Oracle's Spatial extension.

Other advanced features include table inheritance, a rules systems, and database events. Table inheritance puts an object oriented slant on table creation, allowing database designers to *derive* new tables from other tables, treating them as base classes. Even better, PostgreSQL supports both single and multiple inheritance in this manner.

The rules system, also called the *query rewrite system*, allows the database designer to create rules which identify specific operations for a given table or view, and dynamically transform them into alternate operations when they are processed.

The events system is an interprocess communication system in which messages and events can be transmitted between clients using the `LISTEN` and `NOTIFY` commands, allowing both simple peer to peer communication and advanced coordination on database events. Since notifications can be issued from triggers and stored procedures, PostgreSQL clients can monitor database events such as table updates, inserts, or deletes as they happen.

Highly Customizable

PostgreSQL runs stored procedures in more than a dozen programming languages, including Java, Perl, Python, Ruby, Tcl, C/C++, and its own PL/pgSQL, which is similar to Oracle's PL/SQL. Included with its standard function library are hundreds of built-in functions that range from basic math and string operations to cryptography and Oracle compatibility. Triggers and stored procedures can be written in C and loaded into the database as a library, allowing great flexibility in extending its capabilities. Similarly, PostgreSQL includes a framework that allows developers to define and create their own custom data types along with supporting functions and operators that define their behavior. As a result, a host of advanced data types have been created that range from geometric and spatial primitives to network addresses to even ISBN/ISSN (International Standard Book Number/International Standard Serial Number) data types, all of which can be optionally added to the system.

Just as there are many procedure languages supported by PostgreSQL, there are also many library interfaces as well, allowing various languages both compiled and interpreted to interface with PostgreSQL. There are interfaces for Java (JDBC), ODBC, Perl, Python, Ruby, C, C++, PHP, Lisp, Scheme, and Qt just to name a few.

Best of all, PostgreSQL's source code is available under a liberal open source license: the PostgreSQL License. This license gives you the freedom to use, modify and distribute PostgreSQL in any form you like, open or closed source. Any modifications, enhancements, or changes you make are yours to do with as you please. As such, PostgreSQL is not only a powerful database system capable of running the enterprise, it is a development platform upon which to develop in-house, web, or commercial software products that require a capable RDBMS.

Appendix VI - IHE Patterns [15] and HL7 Standards [16]

Due to different HL7 versions present in the different Health facilities involved in the project, the framework will be compliant with HL7 from 2.3 to 3.0. We will provide to the HIS all the reports of the clinical actions possibly performed with our system using the CDA Release 2 HL7 Standard (Clinical Reports). In a similar way both the LIS (Laboratory Information System) and the Anatomic Pathology System will provide to our system the patient's specific data using an HL7 communication standard and following the IHE communication patterns in order to either avoid errors and avoid the manual data entry in the Clinical Data System. Imaging data from the Image Analysis Tools (WP4) and Genomic data coming from the qRT-PCR instrument in HL7 messaging format will be parsed and fully stored in the PRR System.

Here follows a list of documents for the detail of each aspect considered

IHE Cross-Enterprise Document Sharing (XDS) Integration Profile

IHE Cross-Enterprise Document Workflow (XDW) Rev. 2.2-August 31, 2012; IHE IT Infrastructure Technical Framework Volume 1 (ITI TF-1) Integration Profiles Revision 9.0-Final Text -August 31, 2012; Cross-Enterprise Document Media Interchange (XDM), Cross-Enterprise Document Reliable Interchange (XDR), Cross-Enterprise Document Sharing (XDS.b), Cross-Enterprise Sharing of Scanned Documents (XDS-SD)

IHE Anatomic Pathology Technical Framework

IHE Anatomic Pathology Technical Framework Supplement-Anatomic Pathology Structured Reports (APSR)-Trial Implementation-March 31, 2011

IHE Laboratory Technical Framework

IHE Laboratory (LAB) Technical Framework Volume 1 (LAB TF-1)-Integration Profiles-Revision 4.0-Final Text October 2, 2012

HL7 CDA rel.2

CDA Release 2 HL7 Standard (Clinical Reports)

HL7 Clinical Laboratory Automation

HL7 Version 2.5.1 Implementation Guide: Orders and Observations; Interoperable Laboratory Result Reporting to EHR, Release 1

HL7 Version 2.7 Standard: Chapter 13 - Clinical Laboratory Automation

HL7 Version 3 Standard: Laboratory; Result

HL7 Genomic Data Exchange

HL7 IG CG_GENO, R1 Version 3 Genotype, Release1 - January 2009, HL7 IG LOINCGENVA, R1 Version 2 Implementation Guide: Clinical Genomics; Fully LOINC-Qualified Genetic Variation Model, Release 1

HL7 RIM - ISO/HL7 21731

The Reference Information Model (RIM) is the cornerstone of the HL7 Version 3 development process and an essential part of the HL7 V3 development methodology. RIM expresses the data content needed in a

specific clinical or administrative context and provides an explicit representation of the semantic and lexical connections that exist between the information carried in the fields of HL7 messages. The RIM is essential to increase precision and reduce implementation costs. Models are available.

LOINC

Logical Observation Identifiers Names and Codes (LOINC) [17] is a database and universal standard for identifying medical laboratory observations. It was developed and is maintained by the Regenstrief Institute, a US non-profit medical research organization, in 1994. LOINC was created in response to the demand for an electronic database for clinical care and management and is publicly available at no cost.

It is endorsed by the American Clinical Laboratory Association and the College of American Pathologists. Since its inception, the database has expanded to include not just medical and laboratory code names, but also: nursing diagnosis, nursing interventions, outcomes classification, and patient care data set.

LOINC applies universal code names and identifiers to medical terminology related to electronic health records. The purpose is to assist in the electronic exchange and gathering of clinical results (such as laboratory tests, clinical observations, outcomes management and research). LOINC has two main parts: laboratory LOINC and clinical LOINC. Clinical LOINC contains a subdomain of Document Ontology which captures types of clinical reports and documents.

Several standards, such as **IHE** or **HL7**, use LOINC to electronically transfer results from different reporting systems to the appropriate healthcare networks. However, the health information enclosed is identified by a multiplicity of code values that may vary according to the entity producing those results. This has obvious disadvantages to the healthcare network that may need to adopt different codes to access and manage information coming from multiple sources. Managed care providers, for example, often have negotiated contracts that reimburse episodes of care and unique coding to trigger automated claim payment. Mapping each entity-specific code to its corresponding universal code can represent a significant investment of both human and financial capital.

A universal code system will enable facilities and departments across the world to receive and send results from their areas for comparison and consultation and may contribute toward a larger public health initiative of improving clinical outcomes and quality of care.

LOINC is one of the standards for use in U.S. Federal Government systems for the electronic exchange of clinical health information. In 1999, it was identified by the HL7 Standards Development Organization as a preferred code set for laboratory test names in transactions between health care facilities, laboratories, laboratory testing devices, and public health authorities.

Appendix VII - DICOM Standards [18]

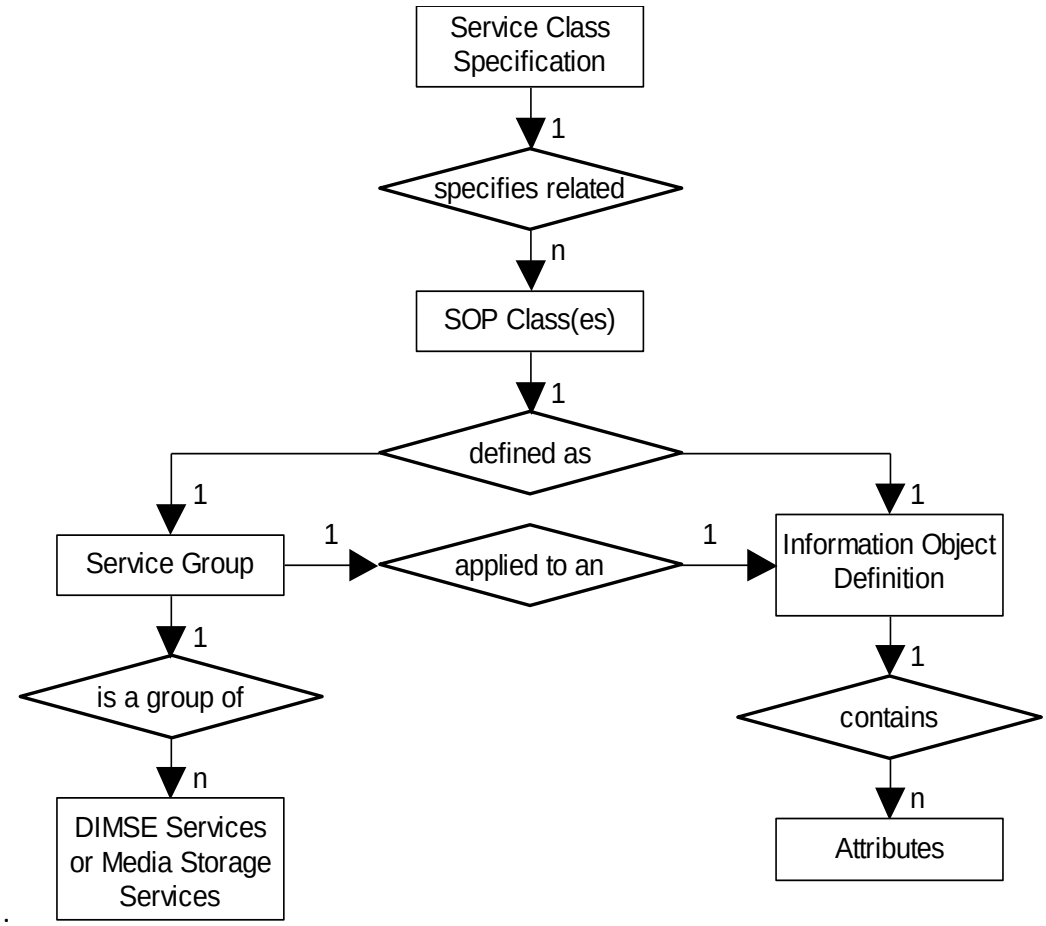


Fig. 6 - Major structures of DICOM Information Model

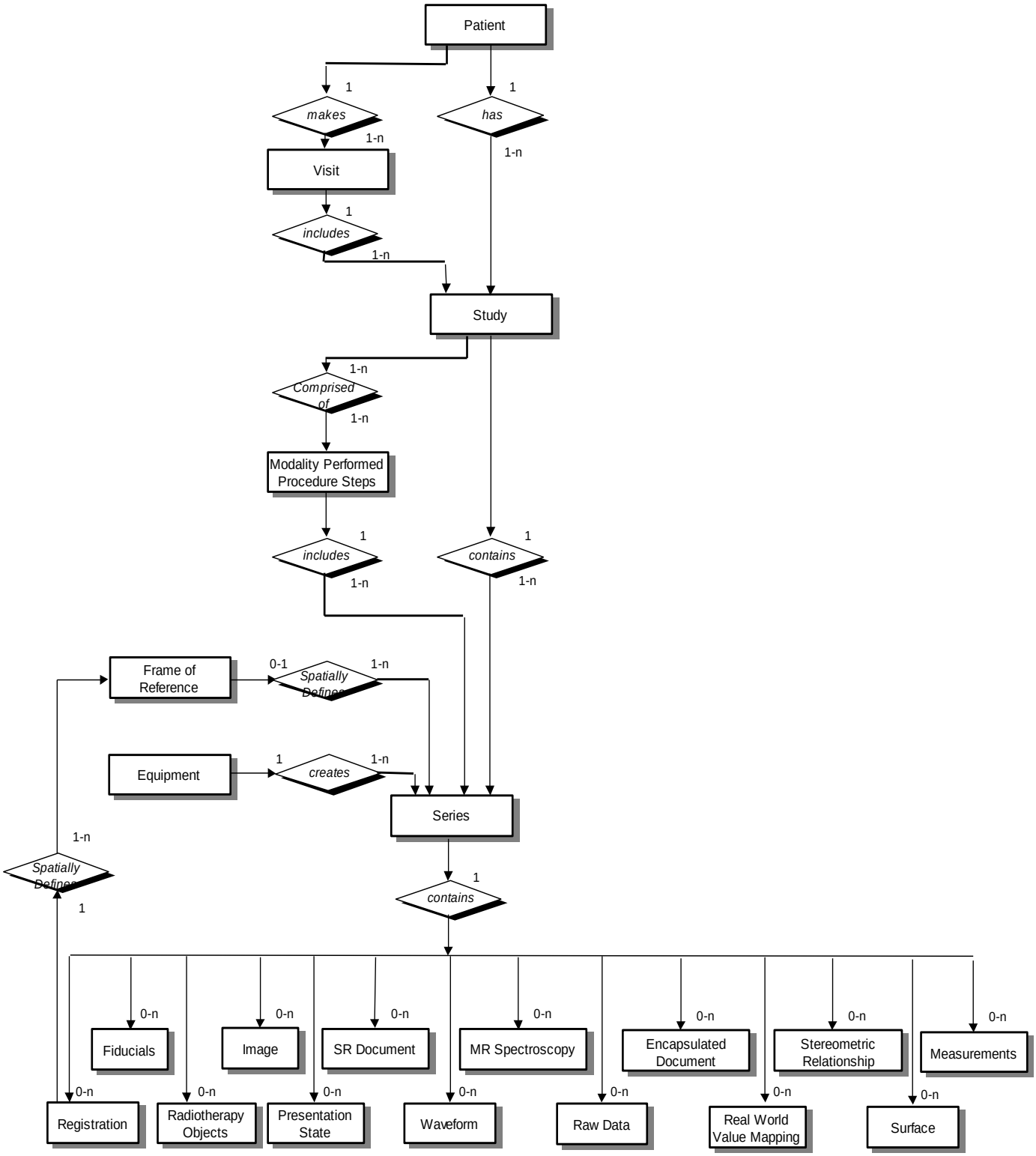


Fig. 7 - DICOM model of the real world

DICOM Query/Retrieve:

This enables a workstation to find lists of images or other such objects and then retrieve them from a picture archiving and communication system.

DICOM Off-line Media:

The off-line media files correspond to Part 10 of the DICOM standard. It describes how to store medical imaging information on removable media. Except for the data set containing, for example, an image and demography, it's also mandatory to include the File Meta Information.

DICOM restricts the filenames on DICOM media to 8 characters (some systems wrongly use 8.3, but this does not conform to the standard). No information must be extracted from these names (PS3.10 Section 6.2.3.2). This is a common source of problems with media created by developers who did not read the specifications carefully. This is a historical requirement to maintain compatibility with older existing systems. It also mandates the presence of a media directory, the DICOMDIR file, which provides index and summary information for all the DICOM files on the media. The DICOMDIR information provides substantially greater information about each file than any filename could, so there is less need for meaningful file names.

DICOM files typically have a .dcm file extension if they are not part of a DICOM media (which requires them to be without extension).

The MIME type for DICOM files is defined by RFC 3240 as application/dicom.

The Uniform Type Identifier type for DICOM files is org.nema.dicom.

DICOM Store & Storage commitment:

The DICOM Store service is used to send images or other persistent objects (structured reports, etc.) to a Picture Archiving and Communication System (PACS) or workstation.

The DICOM storage commitment service is used to confirm that an image has been permanently stored by a device (either on redundant disks or on backup media, e.g. burnt to a CD). The Service Class User (SCU: similar to a client), a modality or workstation, etc., uses the confirmation from the Service Class Provider (SCP: similar to a server), an archive station for instance, to make sure that it is safe to delete the images locally.

DICOM Modality Worklist:

This enables a piece of imaging equipment (a modality) to obtain details of patients and scheduled examinations electronically, avoiding the need to type such information multiple times (and the mistakes caused by retyping)

Appendix VIII - Medical Software Security Standards

While the European health regulatory framework has not a specific directive on the Health IT but still relies on the Directive 93/42EEC about the generic Medical Devices, in 2012, US Congress, promulgated a Federal Law requiring that Food and Drug Administration (FDA) in consultation with the Office of the National Coordinator for Health Information Technology (ONC) and the Federal Communications Commission (FCC) periodically publish "a report that contains a proposed strategy and recommendations on an appropriate, risk-based regulatory framework pertaining to health information technology, including mobile medical applications, that promotes innovation, protects patient safety, and avoids regulatory duplication", the FDASIA Health IT Report [3]. We will use this document as the main reference for a complete regulatory guideline for our project.

In addition to this more wide perspective document we will keep in consideration the following FDA main directives.

Mobile Medical Applications:

”Mobile Medical Applications Guidance for Industry and Food and Drug Administration Staff” published on September 25, 2013. The software environment we will deploy will be under the category of “Mobile apps for which the FDA intends to exercise enforcement discretion”. This category will be applied “For many mobile apps that meet the regulatory definition of a “device” but pose minimal risk to patients and consumers”, the FDA will exercise enforcement discretions and will not expect manufacturers to submit premarket review applications or to register and list their apps with the FDA. This includes mobile medical apps that:

- Help patients/users self-manage their disease or condition without providing specific treatment suggestions;
- Provide patients with simple tools to organize and track their health information;
- Provide easy access to information related to health conditions or treatments;
- Help patients document, show or communicate potential medical conditions to health care providers;
- Automate simple tasks for health care providers; or
- Enable patients or providers to interact with Personal Health Records (PHR) or Electronic Health Record (EHR) systems.

Connected Health

Refers to electronic methods of health care delivery that allow users to deliver and receive care outside of traditional health care settings. Examples include mobile medical apps, medical device data systems, software, and wireless technology. On August 6, 2013, the FDA issued a federal register notice Modifications to the List of Recognized Standards, Recognition List Number: 032 that recognizes voluntary consensus standards to help support and strengthen the interoperability and cybersecurity of networked and connected medical devices.

Cybersecurity

Medical devices, like other computer systems, can be vulnerable to security breaches, potentially impacting the safety and effectiveness of the device. This vulnerability increases as medical devices are increasingly “connected” to the Internet, hospital networks, and to other medical devices. To mitigate and manage

cybersecurity threats, the FDA recommends that medical device manufacturers and health care facilities take steps to assure that appropriate safeguards are in place to reduce the risk of failure due to cybersecurity threats, which could be caused by the introduction of malware into the medical equipment or unauthorized access to configuration settings in medical devices and hospital networks (June 13, 2013, Safety Communication, Cybersecurity for Medical Devices and Hospital Networks) [4].

Confidentiality of data

Confidentiality Policy: as a general rule the whole system and system components will adopt the ISO/IEC 27002 standards in terms of security policy, organizing information security, asset management, personnel security, physical and environmental security, communication and operations management security, access control security, system development and maintenance or system acquisition, incident handling, business continuity management, compliance.

Coming to a more detailed analysis, since healthcare data handling needs the most rigorous approach, it's focal to adopt strict policies on 3 critical items:

1) User Authentication – Mobile devices SIMs are an easy to use highly available Strong authentication support that we will use for both mobile and web interfaces.

2) Data transport encryption – For data transport encryption we will use a 128bit TLS-SSL protocol encapsulating the application-specific https protocol thus securing both the WiFi, Internet and GSM 3G/4G communication from mobile devices.

3) Database encryption – Even if the demographic data, in our architecture, will be drastically separated by clinical data, database encryption is required by some country laws about (e.g. Italian DLgs 30 giugno 2003 n°196).

All those policies will strictly follow the EU Data Protection Directive 95/46/EC as well as any subsequent possible improvement (European Commission - MEMO/14/60 - 27/01/2014) EU will define; in one word we will fully adopt the "privacy by design" concept.