Seirbhíse Sláinte
Health Service Executive

ERS
Electronic Record Services BV

HEALTH SERVICE EXECUTIVE

ICT National Integrated Services Framework Project

Mini Tender: Information Architecture

(work components 1 - 4)

BACKGROUND DOCUMENT

version 5.0

Table of Contents

1. Purpose of this document 4
2. Description of work 5
3. Introduction: Back ground 9
Introduction: Why is background information important 9
Concept: Interoperability 9
Concept: Logical Model of the EHR 13
Concepts: Standards and standardisation 14
Concept: What is the 'Semantic Stack' 17
Concept: Semantic Interoperability Artefacts: Structure and Codes 18
Concept: Two Level Model Paradigm 20
Introduction: State of the Art developments 22
Introduction: Ireland - Workshops and questionnaire - Summary 28
4. Information Architecture Reference Model 29
IA-RM: Introduction 29
IA-RM: Why is it needed 31
IA-RM: Requirements 31
IA-RM: Possible Solution Paradigms 32
IA-RM: Description 45
IA-RM and stakeholders 47
IA-RM: Summary 49
5. Subject Area Model 51
SAM: Introduction 51
SAM: Why are SAM's important 51
SAM: International Developments 53
SAM: Description 55
SAM: Summary 56
6. Determination of potential participating technical systems 57
Technical Systems: Introduction 57
Technical Systems: What is the importance 57
Technical Systems: Summary 58
7. Tools supporting the Subject Area Model 59
Tool: Introduction 59
Tool: Why are they important 60
Tool: SAM Artefact Editor 60
Tool: Document Manager content 62
Tool: SAM Data Dictionary 64
Tool: Terminology services 64
Tool: Technical Testing Service 66
Tool: Summary 66
8. Deployment / implementation 68
Deployment: Introduction 68
Deployment: Why is deployment important 68
Deployment: Environment 68
Deployment: Summary 69
9. Governance: Framework and Tooling 71
Governance: Introduction 71
Governance: Why is governance important 71
Governance: Description 71
Governance: Summary 74
10. Catalogue of deployed Standards 76
11. Introduction: epSOS 81
epSOS developments 82
epSOS: Infrastructure 83

SAM: epSOS dataset 83
SAM: Coding systems as used by epSOS 83
SAM: Semantic Interoperability Artefacts Modeling Method 88
SAM: Semantic Interoperability: Artefacts and Coding Systems 92
SAM: SNOMED coding systems as used 99
12. Appendix: Glossary 101
13. Appendix: Questionnaire 103

1. Purpose of this document

This document is the collection of notes and other background materials that were used to create the synthesis of the work as part of the Executive Summary, Overview and Recommendations document on the Information Architecture Lot (work components 1-4) of HSE ICT National Integrated Services Framework Project.

This background document is produced only for internal use by the Project Steering Board of the HSE ICT National Integrated Services Framework Project.

2. Description of work

The ERS proposed Description of Work in this Mini Tender is summarised in the table. It is indicated where parts of this DoW are covered.

Part	Description	Reference
1.1	**The establishment of a suitable standards based Information Architecture Reference Model (IA-RM)** ERS will perform desktop research on Use Cases and requirements for an IA-RM. ERS will produce a first draft document on the Uses Cases and requirements and how as described it can be mapped to a selection of relevant open standards; meaning relevant for the IA-RM. Stakeholder groups' representatives will be selected by HSE. In minimally 2, maximally 3 sessions, with the stakeholder groups the draft document is discussed and fine-tuned for each stakeholder group. ERS will send a pre-final version of the document to HSE for acceptance.	Chapter 8
1.2	**The provision of a standards based subject area model (SAM)** ERS produces a draft document, based on the IA-RM document, on a Subject Area Model (SAM). Stakeholder groups' representatives will be selected by HSE. In minimally 2, maximally 3 sessions, with the stakeholder groups the draft document is discussed and fine-tuned for each stakeholder group. ERS will send a pre-final version of the document to HSE for acceptance.	Chapter 9
1.3	**The determination of which technical systems should participate in the subject area model (SAM)** HSE will define the list of 'core systems' that need to be evaluated; maximum 2 systems for 2 domains (Medication and Laboratory). ERS will use the validated questionnaires as received to validate the proposed model and investigates the viability of possible integration. Stakeholder groups' .representatives will be selected by HSE. In minimally 2, maximally 3 sessions, with the stakeholder groups the draft document is discussed and fine-tuned for each stakeholder group. ERS will send a pre-final version of the document to HSE for acceptance.	Chapter 12
1.4	**Recommendation of a Governance Framework and Tooling for Model maintenance and expansion** ERS will produce a draft document for a Governance Framework and Tooling for Model Maintenance and expansion. 2 Stakeholder groups' representatives (those responsible for the data architecture and business users) will be selected by HSE. In minimally 2, maximally 3 sessions, with the stakeholder groups the draft document is discussed and fine-tuned for each stakeholder group. ERS will send a pre-final version of the document to HSE for acceptance.	Chapter 13
1.5	**A catalogue with the Standards and associated sub-sections for the Model** ERS will produce a draft document on Standards used plus guidance on their use Stakeholder groups' representatives will be selected by HSE. In minimally 1, maximally 2 sessions, with the stakeholder groups the draft document is discussed and fine-tuned for each stakeholder group. ERS will send a pre-final version of the document to HSE for acceptance.	

2.1	**The determination of a Standards Based Data Dictionary including the specification of meta data structure, data classes, entities and attributes**. ERS will deploy the ISO/IEC 11179 conformant Data Dictionary service. Stakeholder groups' representatives will be selected by HSE. ERS will demonstrate, explain, the deployed Data Dictionary service in minimally 2, maximally 3 sessions, with the stakeholder groups. ERS will describe how Nodes in semantic interoperability artefacts populate the data elements in the HSE Data Dictionary ERS will send a pre-final version of a report about the demonstrations to HSE for acceptance.	
2.2	**Validation (Proof-of-Concept) of the Data Dictionary** ERS will establish the two datasets to be used in the Validation of the Data Dictionary. ERS will convert into two artefact libraries with bindings to relevant coding systems. ERS will populate the HSE Data Dictionary. ERS will demonstrate the artefacts produced, data elements and their relationships in minimally 2, maximally 3 sessions, with the stakeholder groups as selected by HSE. ERS will send the pre-final Validation Document to HSE for acceptance.	
2.3	**Recommendation of a Governance Framework and Tooling for Dictionary maintenance and expansion** The production of standards based artefacts. The assurance of the integrity and quality of the Data Dictionary. Tools to define and publish reference sources: terminologies, classifications, code-sets, value sets. The Data Dictionary maintenance and publication.	
2.4	**A catalogue with the Standards and associated sub-sections for the Dictionary** ERS will produce a first draft document about the standards used. Stakeholder groups' representatives will be selected by HSE. In minimally 2, maximally 3 sessions, with the stakeholder groups the draft document is discussed and fine-tuned for each stakeholder group. ERS will send a pre-final version of the document to HSE for acceptance.	
3.1	**Provision of a blueprint for the deployment, management and maintenance of the terminology service** ERS will perform desktop research. ERS will produce a first draft document on: tooling, deployment, management and maintenance of terminological services. Stakeholder groups' representatives will be selected by HSE. In minimally 2, maximally 3 sessions, with the stakeholder groups the draft document is discussed and fine-tuned for each stakeholder group. ERS will send a pre-final version of the document to HSE for acceptance.	
3.2	**Validation of the proposed model thought the binding of SNOMED CT concepts to a specified clinical dataset** ERS will produce a draft document on the Governance of terminological services. ERS will produce a draft document on the correct deployment of the applied (SNOMED) codes. Stakeholder groups' representatives will be selected by HSE. In minimally 2, maximally 3 sessions, with the stakeholder groups the draft document is discussed and fine-tuned for each stakeholder group. ERS will send a pre-final version of the document to HSE for acceptance.	
3.3	**A catalogue with the Standards and linkages associated with its operation and maintenance**	

4.1	**What standards based toolsets and support structure are required and available to manage the collective practical outputs and relationships of the information components listed above** Work performed in Parts 1, 2, and 3 will be used as input for desktop research by ERS. ERS will consult relevant international eHealth initiatives about relevant experiences with regards to appropriate quality toolsets and support systems to validate the desktop research. ERS will seek advice from relevant Irish organisations on the inclusion of any internal or external specialist skill-sets and service arrangements. ERS will produce a first draft document that can serve as an outline of the toolset in the HSE Infostructure and required specialist skill sets. Stakeholder Groups' representatives will be selected by HSE. In minimally 1, maximally 2 sessions, with the stakeholder groups the draft document is discussed and fine-tuned for each stakeholder group. ERS will send a pre-final version of the document to HSE for acceptance.	
4.2	**A brief comparison of the products that best meet this need including commercial and open source toolsets** ERS will use relevant parts of Parts 1, 2 and 3 as input for desktop research ERS will consult relevant international eHealth initiatives using the outline produced under Part 4.1. ERS will produce a first draft document for discussion with stakeholder groups Stakeholder Groups' representatives will be selected by HSE. In 1 session, with the stakeholder groups the draft document is discussed and fine-tuned for each stakeholder group. ERS will send a pre-final version of the document to HSE for acceptance.	
4.3	**The recommendation of an established toolset and management approach to facilitate integrated governance of the data model, data dictionary and terminology service** ERS will do desktop research on governance, toolsets and management approaches for governance of the data model, data dictionary and terminology service. ERS will provide a first draft document on a recommendation of an established toolset and management approach for governance of the data model, data dictionary and terminological service Stakeholder Groups' representatives will be selected by HSE. In minimally 1, maximally 2 sessions, with the stakeholder groups the draft document is discussed and fine-tuned for each stakeholder group. ERS will send a pre-final version of the document to HSE for acceptance.	

3. Introduction: Back ground

This chapter is an introduction to several basic generic topics that are essential to understand the next chapters about the proposed choices for an Irish National Infostructure.

Topics that will be presented are:

- Semantic Interoperability
- Logical model of the EHR
- Standards and standardisation
- National INFOstructure / Semantic Stack
- Structures and coding systems
- State of the Art Developments

3.1. Introduction: Why is background information important

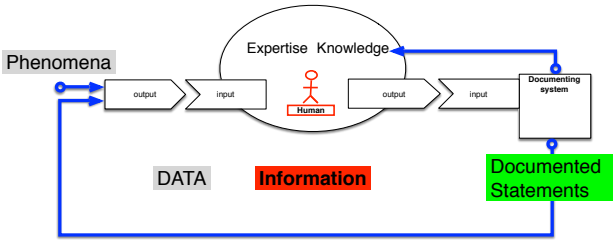
Topics related to semantic interoperability need to be addressed before data can be safely, flexibly, exchanged between different IT-systems.

Reading the topics in this chapter will become clear that:

- Large-scale patient safe semantic interoperability will not be possible without a co-ordinated set of executable policies at a national scale
- an INFOstructure, consisting of several tools for editing, governance and publishing of shared artefacts in libraries, is a conditio-sine-qua-non
- such an national INFOStructure needs to be based on national and international open standardard in order to create a level playing field for the IT-Industry
- the use of open international standards allows cross border interactions between healthcare organisations but also the IT-industry products
- a list of open international standards for the structuring of health information, using codes from coding systems, and local, regional, national, arrangements with respect to datasets are needed
- Subject Area Models (SAM's) leverage/complement existing standards for data exchange as used in Ireland
- The Information Architecture Reference Model and SAMS's fulfill the wish by ISF stakeholders for a seamless integration of data

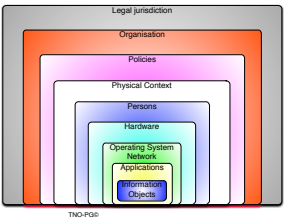
3.2. Concept: Interoperability

Semantic Interoperability is a keyword in this document. This concept will be defined and explained for the purpose of this project.
The general concept 'Semantic Interoperability' can be explained by the phrase: 'the facility, the services, that a healthcare provider can re-use data about a patient now and in the future'.



Many definitions exist each addressing an aspect of the problem stack associated with complex communication between persons, organisations and there IT-systems.
Before discussing semantic interoperability it is necessary to define what is involved in documentation and exchange between communicating parties.
The figure XXX explains the concepts: 'Data', 'Information', 'Documented statements', that all three play their role.
Data is read from a documentation system or data is observed as the result of phenomena that occurred and are perceived and observed. All data is the output of something and read as input by humans. Together with existing implicit and explicit knowledge and expertise the data is interpreted and transformed into 'Information'. After the interpretation of the data inferences are made that result in output such as orders, and explanations, that can be documented in a documenting system. Information is documented. Observe that when read again it has to be interpreted as data again.
In essence semantic interoperability is the feature where Information in one person's brains can be transformed to a signal that when received by another can be interpreted by a third party without a loss of meaning.
Interoperability can take many shapes and forms: between humans, humans and machines, between machines in the form of: sound and other human signals, or via print, digital images, Word files, more and less digital exchange formats.

Complicating factor is the fact that humans are active in work processes and legal systems. Each of these contexts will influence what and how phenomena are observed and interpreted and named. Adding these social and legal contexts make semantic interoperability more complex. In addition there are differences in IT-systems that impact the exchange between communicating parties.



In summary many and different aspects play a role in semantic interoperability.
Interoperability is a property referring to the ability of diverse systems and organizations to work together (inter-operate).
The IEEE defines interoperability as the ability of two or more systems or components to exchange information and to use the information that has been exchanged.

What is needed for full patient safe Semantic Interoperability?

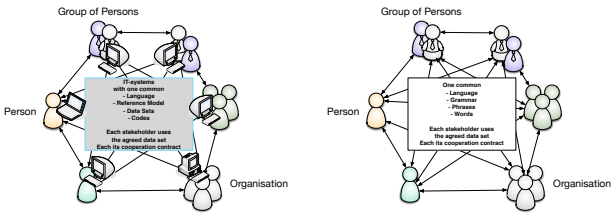
Technical interoperability
In order to convey a message from one communicating party to another a mutually agreed medium is needed for the transport. E.g. air for sound, paper and the postal system for written letters, or electronic messages via the e-mail system on the Internet or another network.
This level needs standards for Common Transport Models.

Syntactical interoperability
For this level of sophistication agreements between the communicating parties are necessary about formats such that the texts can be read and interpreted. For example both must agree on a Word, or PDF, or image format, or XML formats that can be used to hold the message that is to be exchanged. This level of interoperability needs the interpretation by humans to decipher the message. IT-Systems can display and print the message but cannot interpret it and act upon it.
This level needs Common Message Models.

Semantic interoperability
At this level of sophistication the message that is exchanged contains not only human readable text but additions that allow IT-systems to interpret it and act upon it safely.
Two sub-levels can be discerned: Partial Semantic Interoperability and Full Semantic Interoperability.

Partial Semantic Interoperability
For semantic interoperability between humans to work agreements are necessary on the topics of: language to be used, character set that is associated with the language, its grammar (syntax) and the common dictionary with the meanings of the words, plus standard phrases.

When persons, groups of persons or organisations make use of IT-systems, these systems must be able to read and produce similar constructs. E.g.: Language, Reference Model, Datasets, and Codes.

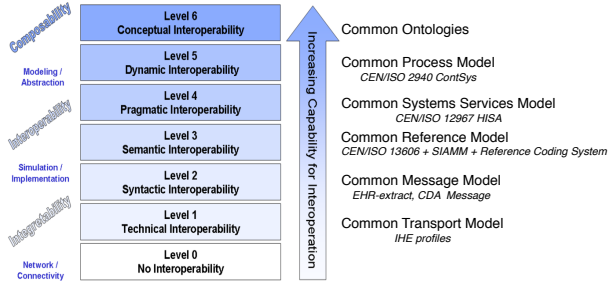


Arrangements on these topics between two or more communicating parties allow humans to interpret the data safely.

When structures and codes are use only partial interoperability is possible, because only humans can understand the text and its associated code. Humans know what 'cold' or ' warm' mean. Healthcare providers know what 'pneumonia' means. A substantial amount of implicit knowledge is necessary to understand fully the meaning of the message. Some implicit meaning is hidden in the software of applications.

Conceptual Interoperability
Full Semantic Interoperability
IT-systems can deal in a limited way with the data using software, since they do not have access to all this implicit encyclopedia knowledge. A full understanding by the IT-system is possible only when the IT-systems have a complete knowledge about the world. For full Semantic Interoperability they must have access to an Ontology as analogue to the encyclopedia.
This level needs Common Ontologies.

The present state of the art has not reached the level of full semantic interoperability. Ontologies and healthcare are considered open for academic studies and experimentation.
Partial semantic interoperability is achievable.



Pragmatic Interoperability
Technical IT-System Co-operability

Persons, groups and organisations make use of IT-systems and shared services. This means that IT-systems of various vendors need to be able to deal with these common services. Examples of services are facilities to look for codes or translations of codes from one language to another, etc.
At present IT-systems have proprietary solutions for these internal and external services. Sharing of services is limited.
This level needs one Common Systems Services Model.

Dynamic Interoperability
Process Co-operability

Each healthcare provider and organisation is using its own work processes. Increasingly there is the need for co-operation between the different stakeholders in clinical pathways. To reach this level of interoperability shared agreements (models) are needed to map the various process steps between organisations and the data collected and shared.
This level needs one Common Process Model.

3.3. Concept: Logical Model of the EHR

The Electronic Health Record (EHR) is a loosely defined concept.
ISO/TR 20514 defined the Shared EHR as: "*a repository of information regarding the health status of a subject of care in computer process able form, stored and transmitted securely, and accessible by multiple authorised users. It has a standardised or commonly agreed logical information model which is independent of EHR systems. Its primary purpose is the support of continuing, efficient and quality integrated health care and it contains information which is retrospective, concurrent, and prospective.*"

For practical purposes this project will define the EHR as those services that allow the co-operation between healthcare providers and their organisations. Each of the connected systems that healthcare providers use - the Electronic Medical Record (EMR) and Personal Health Record (PHR) - are NOT part of this EHR definition.

The logical model for this EHR is depicted in the figure above.

Healthcare Providers use an Electronic Medical Record system (EMR); patients use Personal Health Record system (PHR). Data that is shared is placed, as it were, on an Exchange Plane. Data can be there for a short period of time or forever, depending on the needs. Examples are: Patient Summary, discharge and referral letters, disease specific shared records, registries, etc.
All data in that Exchange Plane are outside the confines of user systems. Access and privacy is controlled via an Access Control List (ACL) and Patient Mandates that the author of the published data together with the patient are responsible and accountable for. ACL's are mostly associated with access to services. At the level of the data in the patient record Patient Mandates govern the access to specific kinds of data. Any needed override of the ACL via the 'Red Knob' procedure needs an immediate notification of the author and patient.
Out of hours substitution can be handled in the same way as exchanges between systems, but actually and logically the original other EMR from the Healthcare Provider they substitute, is used.

Other stakeholders (researchers, registries, authorities) can re-use the data that is made available via the data exchange plane.

IT-systems (EMR, and PHR) that exchange data need to have a common interface based on the same specifications for all the models and coding systems needed for Semantic Interoperability, and as described in the previous chapter. Each of the existing EMR and PHR systems internally can be different as long as at the interface with the Exchange Plane they use the same specifications to normalise the exchanged data.

3.4. Concepts: Standards and standardisation

Any infrastructure for the exchange of health data using the common agreed interface specification between health-IT systems must be based on many standards. This chapter will explain standardisation, the need for it, and the most important standardisation organisations. Elsewhere the standards that are needed are listed and explained.

3.4.a. What is a standard?

A standard (French: Norme, German: Norm) is a *technical document designed to be used as a rule, guideline or definition. It is a consensus-built, repeatable way of doing something.*
Standards are created by bringing together all interested parties such as manufacturers, consumers and regulators of a particular material, product, process or service. All parties benefit from standardisation through increased product safety and quality as well as lower transaction costs and prices.

Standards can be organisation specific (industrial standard), in the private (closed) or public domain (open), national (NSAI, ANSI, DIN, BSI, …) or international, regulated or non-regulated.
Most National projects made the choice to rely on open and public standards. In the case of Europe they by preference use European standards.

Because of European Directives and regulations standards can play its role in procurement processes.

3.4.b. European standardisation

Three European Standards Organisations (ESOs) are officially recognised as competent in the area of voluntary technical standardisation. The European Union (EU) Regulation (1025/2012) which settles the legal framework for standardisation, has been adopted by the European Parliament and by the Council of the EU, and entered into force on 1 January 2013.

The three European Standardisation Organisations are:

- CEN[1] – European Committee for Standardisation;
- CENELEC[2] – European Committee for Electrotechnical Standardisation;
- ETSI[3] – European Telecommunications Standards Institute.
- Thanks to European Standards created by the European Standardisation System, manufacturers and service providers get direct access to the market. The ultimate aim is to have a unique standard in 33 European countries and beyond.

Indeed, by creating standards, the ESOs help facilitate trade between countries, create new markets and cut compliance costs. They provide a Standardisation framework to prepare voluntary standards that help to develop the Single European market for goods and services via procurement. European Standards play a crucial role in the development and consolidation of the European Single Market and support other EU policies, providing an efficient co-regulation tool.

Over the years the European Standardisation System has proved to be a successful and well-consolidated model.

The European Standards Organisations formally cooperate with other international standardisation organisations such as: ISO4, IEEE5, GS1, etc..

[1] www.cen.eu

[2] http://www.cenelec.eu

[3] http://www.etsi.org

[4] http://www.iso.org

[5] http://www.ieee.org

3.4.c. Other International Standardisation Organisations

In the field of healthcare other standardisation organisations play an important role: HL7[6], IHTSDO[7], LOINC[8], etc..

3.4.d. National INFOstructure / Semantic Stack

When in a region or country the various stakeholders need to exchange data about patients a set of shared agreements is necessary.
This set of agreements is about topics such as:

- semantic aspects
  - unique identifiers for objects: patients, healthcare providers, provider organisations, other stakeholders, services, etc.
  - unique codes from relevant coding systems and mappings between them
  - dataset definitions of what is exchanged between stakeholders
- technical exchange formats
- governance of the common and shared agreements.

3.4.e. What is the 'INFOstructure'?

The Infostructure is all that is needed, as a shared infrastructure, in a country or region to make semantic interoperability possible.

[6] http://www.hl7.com

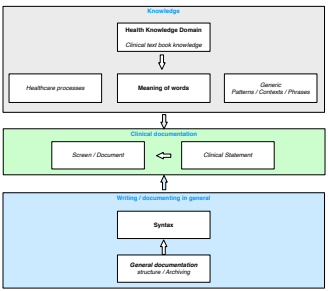
[7] http://www.ihtsdo.org

[8] http://www.loinc.org

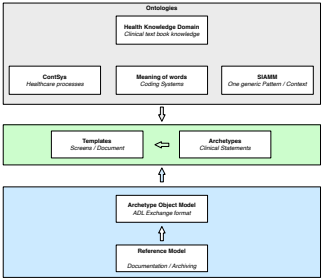
3.5. Concept: What is the 'Semantic Stack'

The Semantic Stack is all that is needed, as part of the INFOStructure, in a country or region to make semantic interoperability possible.

In human written exchange of data several things need to be in place:



- agreements about how to produce and archive documents/screens
- agreements about the language specific syntax used in the document/ screen. (Observe that this allows the construction of syntactical correct nonsense)
- agreements about the context specific layout for a document/screen
- agreements how to construct meaningful sentences to be used as Clinical Statements in a specific context
- agreements about the use of words and their shared meaning
- agreements (implicit or explicit) about the back ground knowledge about the (specific aspects of) the world.
- In computer assisted interoperability the same components are necessary:
- agreements about a Reference Model that deals with the general structure of any document and archiving issues
- agreements a method to produce sentences, a kind of syntax using an Archetype Object Model
- agreements on the structure of specific documents (Templates) representing a specific dataset
- agreements on the construction, structure, of Clinical Statements (SIAMM)
- agreements on the content of Clinical Statements (Archetype)
- agreements on the Meaning of all Words/concepts (Codes and Coding Systems) used in Clinical Statements and Documents



- agreements on the shared specific and general health knowledge (Ontologies)

Only a complete set of agreements, standards, and governing organisations at a regional or national level can realise flexible and safe exchange of patient data between IT-systems in health and care. Only a stack of standards that is used by all stakeholders will be able to create the Exchange Plane as described in the chapter on the Electronic Health Record. This stack of standards is depicted in figure XXX and is called the Semantic Stack.

When IT-systems in an interface now how to deal with data, that is exchanged according to the open International standards of the Semantic Stack, they can stay as they are and use company specific proprietary software inside their systems.

3.6. Concept: Semantic Interoperability Artefacts: Structure and Codes

National or regional or local healthcare actors express their data needs in a dataset. This is what their systems must be able to collect and exchange.
As presented in the previous paragraph elements from the Semantic Stack are used.
This paragraph describes generically the principles behind semantic interoperability artefacts and the crucial role that codes play.

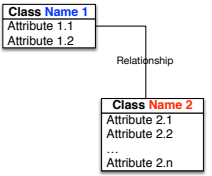
In the case of human communication we rely on many shared constructs (syntax, phrases, words and a shared knowledge about the world.
Equivalents in semantic interoperability between computer systems are:

- Knowledge - Ontologies, defining knowledge
- Words - Codes from coding systems, defining meaning
- Phrases - Archetypes / Clinical Statements, defining structure
- Specific Document - Template, defining specific structure and meaning in a specific context
- Generic Document - Reference Model, defining a general structure

Archetypes are used to construct the specific Templates that represent a report or screen.
Codes are used to give meaning to specific parts of Archetypes.
Concepts from ontologies give meaning to codes.

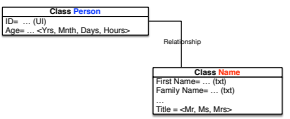
In general communication amounts to agreements on shared Knowledge, Structures and Meaning/Codes.
And when the communication is digitally this Knowledge, Structure and Meaning/Codes need in addition internal technical representations that computers can manipulate.

Archetypes as Structures based on a Reference Model and codes are the main atomic building blocks that are used.



The Reference Model and therefor all archetypes consist of UML classes with one or more Attributes per Class. The UML model allows the definition of the relationships between classes. The UML model provides the possibility to structure data.
Per UML CLASS the name of that class (Node) is specified and in addition each UML CLASS holds one or more Attributes that contain data values.

When modeling archetypes the structure, names of nodes, the attribute names but also the data values attached to these attributes can be specified.
In the example a Class named Person is connected with a sub-class named Name. About the Person and its Name things are specified in attributes that define what data can be captured.

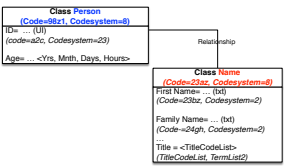


Archetypes are constraints on the Reference Model and therefor always are conformant to this Reference model. Archetypes are built using UML classes that are named. UML Classes have Attributes that are named also. Data values associated with these Class Attributes can be of many kinds: numbers, texts, codes, dates, times, url, unique identifiers, etc.
The names provided are arbitrary and have only meaning to humans. These names can be translated in any language. Inside the archetype and the systems that use it these names are represented by internal codes. Each internal code potentially can have any name. Names that have meaning for the user that provided that name. Each internal code does NOT carry a universally understood meaning of what it exactly is. The meaning attached to these names of Classes, Attributes and data values are derived from codes from one or more Reference Terminologies.

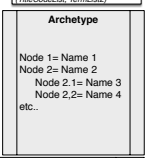
All these names and possibilities that are given to the Node names, Attribute names and data values can be for instance texts, numbers, dates, times, etc. or any combination. Data types specify the format of those data points.
Each author, each community potentially can specify its own structure, Node names, Attribute names and allowed content of the data points. Thereby not securing any general semantic interoperability.

For semantic Interoperability one basic structure (Reference Model), one way to construct and exchange artefacts, a set of common Node names, one set of common Attribute names, one set of common allowed Data types, one set of allowed or expected data values and one set of unique codes for all these names and data values are necessary.

Sometimes the codes are attached during design time. Sometime the codes for (for instance) possible Diagnosis must be selected from an available list of pre-selected codes to choose from at run-time. These lists are called: TermLists. These TermLists are stored outside the EHR and made available as catalogue to choose from.

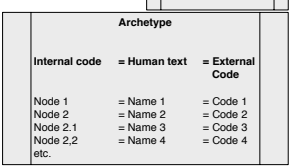


An important benefit of the use of an international coding system is that while codes stay the same, the accompanying definitions and labels can be in any language. International coding systems play an important role in cross border, cross language, semantic interoperability.



Because there must not be too many competing coding systems to choose from, some coding systems are recognised as Reference Terminologies.

The INFOstructure plus tooling and libraries with defined artefacts are needed at a national level in order to secure patient safe semantic interoperability. For cross border exchanges International standards are essential.



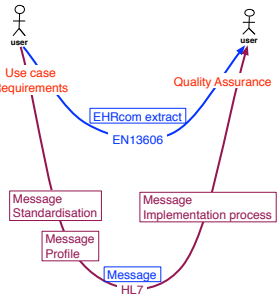
3.7. Concept: Two Level Model Paradigm

The Two Model Paradigm (actually the multi model paradigm) is a relative recent development by CENtc251 and later ISOtc215. It is based on 20 years of European research and resulted in a CEN/ISO standard for EHR-Communication.

Two (or more) models are needed:

- A Reference Model that deals with the structure and archiving of documents
- An Archetype Object Model that allows the creation of Archetypes as constraints on the Reference Model. Each archetype always conforms to the Reference Model. Each archetype can fully specify the data needs of the users.

Based on local data needs by a community they can specify with great detail their local data requirements by grouping archetypes from a library in a structure (named Template). A Template defines for instance the data content of a report or screen or message.



The ubiquitous Message paradigm, as described, needs an implementation process that needs the definition of a message standard per topic, an implementation (IHE[9]) profile per the message standard and an extensive testing process for IT-systems that claim conformance to the message standard.

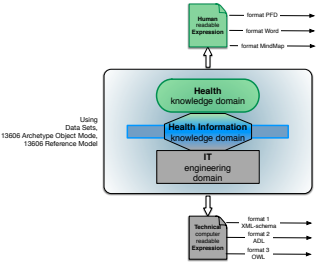
The Two-Level-Modeling paradigm when implemented in a system or when a system has a 13606 based connector can specify ad-hoc EHR-Extracts (Templates) when they are needed as part of a co-operation process that needs exchange of data. Tooling (an archetype editor) creates the EHRcom extracts and allows the creation of technical implementation artefacts that the IT-Industry can implement directly supposing there is the 13606 based connector.

This Two Level Model paradigm has substantial advantages over other paradigms:

- Models drive the archetypes enforcing that every archetype conform to both models
- It allows Model driven development and a flexible fast implementation of EHR-Extracts
- Data bases using this Two Level Model paradigm do not need data base conversions when data requirements change
- Resulting in an agile, flexible, resource friendly solution for semantic interoperability, e.i. seamless integration of data
- Allows a complete separation of concerns between knowledge domain actors (healthcare providers and health organisations) and the IT-providers, while using the same paradigm and artefacts.

[9] http://www.ihe.net

Editor tooling based on open international standards and the Two Level Model paradigm create the Archetypes. Archetypes can be exported such that healthcare providers and their organisations can read, comment and validate the archetypes. While at the same time IT-vendors are provided with technical artefacts based on the same archetype. Technical artefacts such as XML, they can use directly in the development of their systems or use it for integration (semantic interoperability) between systems.



3.8. Introduction: State of the Art developments

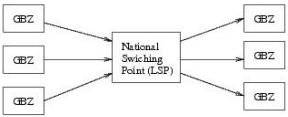
Ireland is not the first or only country that makes plans for a National INFOstructure. Relevant and important recent developments in various countries are described.

3.8.a. Netherlands[10]

The Dutch Electronic Patient Record (EPD) System is a Dutch Nation-wide system for exchanging medical records, which is introduced in 2009-2010. The Dutch senate has decided not to accept a law that regulates and mandates the use of the EPD for exchanging patient information in the Netherlands.

An Association funded by the National Insurers together with Healthcare Provider organisations exploits the developed infrastructure using HL7v3 messages as technological solution.

The EPD is generally characterized as a decentralized system. Patient records are stored in the systems used by the care professional(s) - i.e., the responsibility for managing and storing these records remains with the care professionals; records are not stored in a central database as in, for example, the SPINE system used by the U.K. National Health Service.



[10] Guido van 't Noordeinde (2012) University of Amsterdam - http://staff.science.uva.nl/~noordend/epd/index-start.html

The system's core is the National Switching Point (LSP in Dutch). This system contains a reference index which stores references (pointers) to patient records. Patient records are indexed using a unique identifier for patients (BSN, the former Dutch social security number) and an information type. Access control takes place centrally in the LSP, based on authorization of the care professional for a given information category (e.g., GP record or pharmacy record). The patient records in the EPD are in most cases be professional summaries created by physicians for the purpose of sharing information with colleagues.

The de-central information systems that care professionals store their records in and which are connected to the LSP are termed well-managed care systems (GBZ systems). Only systems that adhere to the requirements for GBZ systems can connect to the LSP.

Above, a figure showing the LSP in relation to GBZ systems is shown. The central role of the LSP is clearly visible. To the right, GBZ systems (belonging to different organizations) are shown which registered patient information in the LSP. Clients (physicians or mandated employees in a GBZ, left) can access the central reference index in the LSP to find relevant records, or they can construct a query to let the LSP find and retrieve relevant records. All access is mediated by the LSP. In reality, GBZ systems will contain client as well as server functionality.

GBZ systems may be small (e.g., GP systems) or very large - including hospitals containing many different systems that contribute information to the EPD, or from which requests are made.

3.8.b. England

A few years ago, the NHS (National Health Service) launched a programme for centralising connected health in the UK called “Connecting for Health”. One of the first initiatives arising from this project was the creation of Summary Care Records (SCRs), the aim of which was to provide healthcare staff treating patients in an emergency with faster access to key clinical information. The NHS also set up HealthSpace, an online service whereby healthcare professionals could share and access medical data. The portal was however not as popular as expected and was closed down by the Department of Health in early 2013.

The NHS in England for many years maintains a data dictionary based on a subset of EN ISO 13606 and SNOMED codes using the Logical Record Architecture[11] (LRA).

As of November 2013 a new phase has started. The Information Standards Board for Health and Social Care now encompasses Department of Health and NHS England representatives.
The health and Social Care Act (2012) states that the following must have regard to an Information Standard published under the Act: the Secretary of State for Health, the NHS, Public Bodies with health services or adult social care, and Anyone providing publicly funded health services or adult social care commissioned by or on behalf of a public body.

[11] http://www.uktcregistration.nss.cfh.nhs.uk/trud3/user/guest/group/0/pack/12

The development of Information standards is guided by the following principles [12]:

- There must be a clear national mandate
- Professionally endorsed and universally applied information standards are important to the delivery of care, especially where information is to be shared. They should support innovation, improved outcomes, comparability and efficiency.
- Standards for record keeping need to be led and endorsed by professional bodies, with appropriate cross-disciplinary governance, and where appropriate incorporated into professional accreditation.
- Standards have a lifecycle from inception to retirement. They should be removed when they have exhausted their utility, and they must not compete inappropriately with each other.
- Information should be captured once, at the point of care. This ensures high levels of data quality and reduces the burden of information collection on the NHS.
- Confidentiality, security and clinical safety must be paramount when using information systems. The ISB provides standards in each area.
- International standards should be used in preference to national and local ones. Healthcare is not just limited to England. Using international standards allows information to flow across borders and reduces the amount of tailoring required when buying international IT systems.

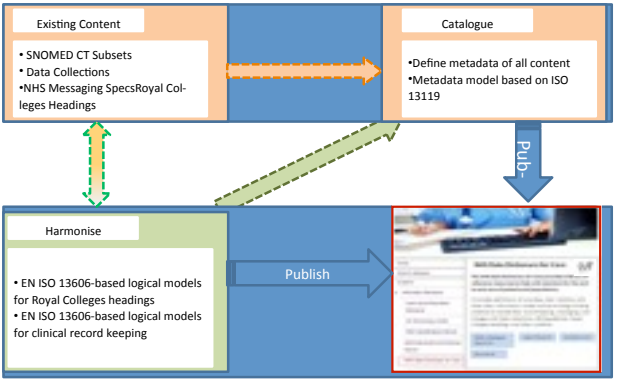
The Information Standards Delivery department provides the building blocks:

- Terminologies: Read, SNOMED-CT, dm+d[13])
- Classifications: OPCS-4[14], ICD-10
- NHS Data Dictionary: Data collections, Datasets
- NHS Data Dictionary for care (DD4C): Data recording/ record keeping, Terminology bindings

[12] Information Standards Board - http://www.isb.nhs.uk/setting/hscact2012/index_html

[13] http://dmd.medicines.org.uk/DesktopDefault.aspx

[14] OPCS Classification of Interventions and Procedures version 4 - http://en.wikipedia.org/wiki/OPCS-4



The EN ISO 13606 EHR-communication provides the foundation to define datasets.
The DD4C is based on the ISO 13119 Standard for Clinical knowledge resources - Metadata.

STEP[15] (Standards Enforcement in Procurement) is in use to create the level playing field for all vendors.

It is the aim of this new development to:

- Provide clinically assured, quality assessed, process-driven logical representations of health care records
- Single reference point for all product-dependent modelling work such as NHS Data Dictionary and NHS Messaging Specifications
- Provide metadata for our logical models to provide valuable information about the models as well as to allow associations with other content such as SNOMED CT subsets, message templates etc.
- Allow multiple format download of our logical models as a free public resource: ADL, UML, HTML, XML, JSON, Mindmap, Word etc.

[15] STEP - http://www.isb.nhs.uk/use/step/index_html

3.8.c. Australia

The Australian National eHealth Transition Authority (NEHTA[16]) is the Austrian governmental initiative that is jointly funded by the Australian Government and all State and Territory Governments. Its goal is described in a strategic plan[17].

NEHTA's Mission:

NEHTA is the lead organisation supporting the national vision for eHealth in Australia; working openly, constructively and collaboratively with consumers, healthcare providers, funders, policy makers and the broader healthcare industry; to enable safer, higher quality, secure, accessible, equitable, efficient and sustainable healthcare.

Key features are that they work on the definition of HL7 Message standards to be implemented by all stakeholders, they will rely on SNOMED-CT-AU and provide testing facilities for software vendors that claim conformance.

3.8.d. New Zealand

Population is 4.5 million. The National Government has developed a National eHealth Strategy and Plan. Its goal is to achieve high quality healthcare and improve patient safety and by 2014 have a core set of personal health information available.

Technical aspects
National Patient, Healthcare Provider and Facility Identifiers are available.
Demographics information is stored centrally (name, address, date of birth, sex and ethnicity).

Standards used will rely on:

- IHE XDS messaging and document handling info structure
- Predefined semantic building blocks as extensible Core Health Content Models (archetypes) using openEHR (an EN ISO 13606 related proprietary specification)
- CCR, HL7 v3 CDA documents, HL7v2, HL7v3.
- ISO 13606 (EHR), 11179 (Data Elements/Dictionary) , 21090 (Data Types)
- SNOMED-CT, LOINC, ICD10, HSSP CTS2

3.8.e. Canada Infoway[18]

The last years Canada has gained much experience with health IT.
Canada is a federal state with a distribution of legal power between the National Government, ten Provinces and three Northern Territories. The Federal Government sets the framework and contributes to the financing arrangements are autonomous. Because of this regional differences are quite substantial.

---

[16] Health * Soxial http://www.nehta.gov.au/about-us

[17] NEHTA Strategic Plan (2011–2012) PDF (

[18] https://www.infoway-inforoute.ca

EHR
In March 2001 Canada Health Infoway was started as an independent organisation owned by vice ministers for health of the fourteen Federal, Provincial and Territories. Infoway provides financial support for a common EHR framework. Best practices are shared. Infoway cooperates with various (local) governments, health organisations and IT-suppliers.

Projects
From 2001 more than 2200 projects have been subsidised: EHR, exchange of Lab reports, tele-health, and patient access to the EHR data. In 2006-2007 350 million Euro was spent. (Canada has 32 million inhabitants.

Problems
The public at large does not support the Canadian vision. A lot of exchanges take place using paper. Support by local governments could be improved.

Technology aspects
Developments in Canada are based on the messaging paradigm (HL7 v3 messages) and SNOMED-CT, LOINC.
Development of messages is supported by 'the Message Builder', Testing Environment, Terminology tools and tools for the localisation of the messages.

Plans for 2015
Deployment of the EHR nation wide, improved public support, better disease management, better tele-care, treatments of cancer patients and a reduction in waiting lists.

3.8.f. European Commission + Member States

epSOS is a European e-Health project executed by almost all (23) EU-Member States for the cross border exchange of digital exchange of relevant patient data for mobile citizens. The project is started in 2009 and is called a Large Scale Project (LSP).
epSOS develops solutions that do not need changes in national legal arrangements. All exchanged data is translated into the language used in the receiving country.

Architecture/Technology
Based on the message paradigm using HL3 v3 CDA, Integrating the Health Enterprise (IHE) profiles, SNOMED and LOINC messages have been developed for the Patient Summary, eMedications and e-Prescribing. Services developed are: Identification Service, Patient Service, Order Service, eDispensation Service and the Consent Service.
Each country has one responsible National Contact Point that links with National users and other National Contact Points. Languages and codes are translated using the Taxonomy Manager and the Terminology Serviced Access Manager, when data is exchanged between the contact points.
In addition there are Auditing and Authentication services.

Status

In 2013 in a few countries the exchange was deployed and tested in a Proof of Concept setting. epSOS is in the process to evaluate the technology choices. The CDA documents (messages) developed at present are in use in the collaboration projects between the EU and USA.

3.9. Introduction: Ireland - Workshops and questionnaire - Summary

[To be added]

The full text of the questionnaire will be included

Alleen de uitkomsten van de vragen; verder geen namen of vrije losse opmerkingen tekst.

4. Information Architecture Reference Model

This chapter corresponds to part 1.1 of the project deliverables.
Based on workshops, the questionnaire, use cases and requirements an analysis was performed.
It was decided that the epSOS patient dataset was to be used as the use case.
This approach was validated by the questionnaire and the workshops.

4.1. IA-RM: Introduction

The Information Architecture Reference Model (IA-RM) is a high level abstract model that depicts how at the level of exchange of health data (messages, documents and objects, plus coding systems) Semantic Interoperability Artefacts (Outputs) are produced, maintained and published, including the set of supporting services.
The outputs reflect the datasets healthcare providers produce and validate. A collection of outputs for a specific domain are called Subject Area Models (SAM's). The proposed HSE Tooling Environment will support all stakeholders to define SAM's in a coherent governed way.

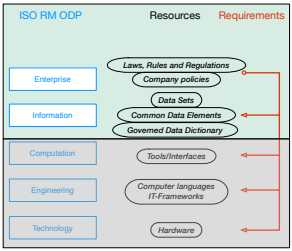
The ISO RM/ODP[19] standard provides a generic framework to describe the various aspects of IT-systems. It discerns 5 levels, called Viewpoints:

- The **enterprise viewpoint**: A viewpoint on the system and its environment that focuses on the purpose, scope and policies for the system.
- The **information viewpoint**: A viewpoint on the system and its environment that focuses on the semantics of the information and information processing performed.
- The **computational viewpoint**: A viewpoint on the system and its environment that enables distribution through functional decomposition of the system into objects which interact at interfaces.
- The **engineering viewpoint**: A viewpoint on the system and its environment that focuses on the mechanisms and functions required to support distributed interaction between objects in the system.
- The **technology viewpoint**: A viewpoint on the system and its environment that focuses on the choice of technology in that system.

[19] ISO/IEC 10746 'Open Distributed Processing — Reference model'

This document will focus on:

- the requirements from the environment, healthcare providers, etc. as expressed at the enterprise viewpoint level.
- the semantics (SAM's) that express the data needs healthcare providers, their organisations and authorities have, as expressed at the information viewpoint level.
- the behavior of IT-systems, applications and services, in casu the HSE Toolbox, will be described in the computational viewpoint.
- Engineering and technology requirements are out of scope for this document.



4.2. IA-RM: Why is it needed

As long as each user in one IT-system uses its own software and locally defined codes, the internal interoperability operability is 100%. The moment they start to communicate with other 'autonomous' IT-systems problems arise. Each IT-system has their own data base structure, define concepts and codes. They can not communicate. There is NO semantic interoperability.
An Information Architecture is needed. The Reference Model describes the architecture.

A solution that has been tried is to select for one region or one country one commercial vendor that provides the EHR functionality. This solution has several serious drawbacks:

- One vendor means automatically a vendor lock-in. Changing form one proprietary solution to an other is complicated and will lead to patient unsafe data loss.
- One vendor for the national/regional EHR solution means that the local IT-vendor market will change. There is a serious risk that the local, national, vendor market will be damaged beyond repair. (E.g. the effect of a few vendors that were allowed in the market place in England)

One other possibility is an IA-RM as a collection of semantic interoperability artefacts (messages) with codes, as is the case in several countries (e.g. the Netherlands where an IA-RM is constructed by defining a 'static' set of HL7v3 messages).The preferred solution is one Information Architecture Reference Model that normalises as much as possible all data that is exchanged. This solution secures a thriving national, regional, IT market for EHR solutions and prevents vendor lock-in.

The IA-RM consists of a library of modules (based on a set of requirements) of re-usable semantic interoperability artefacts and supporting tooling that can be used flexibly to construct any message needed for reporting or shared care. This IA-RM is used to define datasets/SAM's and create SAM based exchange formats for communication between IT-systems as part of a procurement process for IT-systems and supporting tooling and services. All this is called the INFOstructure.

An IA-RM is useful when healthcare providers and their organisations need to report all to regional or national or international shared registries. An IA-RM is useful, also, when different health actors need to co-operate around the shared, joint, delivery of health and care to the same patient.
An IA-RM plus supporting tooling allows each user community to re-use building blocks for creating Subject Area Models (SAM's) that reflect a dataset that is needed locally or prescribed by law.

4.3. IA-RM: Requirements

In order to make choices that lead up to one National solution for Semantic Interoperability requirements need to be collected.
Based on meetings, discussions, a questionnaire, and desktop research the following list of requirements is used.

As a pre-amble, New Zealand[20] published a set of general principles that in adopted form are relevant. The following high-level principles underpin the Information Architecture Reference Model for Ireland and will guide its development:

Align to national strategy. The Reference Model for the Information Architecture will align with national standards and business strategies, with priorities defined by national IT plans.

- Invest in Information. We will represent health data for exchange as detailed clinical models that can be represented in different ways independently of any particular information model or serialized representation (structure) and derived directly from business requirements with clinical input. These models may be represented in different ways for different audiences.
- Use single content model. Information for exchange will be defined and represented in a single consistent way at the information model level. Where possible, it will align with national and international standards.
- Work with sector. The development of the Reference Model for the Information Architecture will be in partnership with the sector as represented core groups as relevant Irish stakeholders, such as: (1) Business and Strategy; (2) Clinical, Safety and Research; (3) Technical; (4) Allied Agencies of HSE and (5) Others, like industry, etc. …
- Align to business needs. Development of the details of the Reference Model for the Information Architecture will be in conjunction with the prioritized business projects. Prioritization will be set by health and IT plans embodying those needs. The intent is to ensure clinical and other business engagement.
- Use proven standards. Where there is a relevant national or international standard that is compliant with the overall direction of the Reference Model for the Information Architecture, will meet a particular business/technology requirement and is used, we will use that standard. If modifications are required, we will work with the relevant SDO to make the modifications. This approach applies at all levels of the interoperability stack including workflow, payload, security, terminology and transport.
- Adopt services approach. To define the behavioral aspects of interoperability we will use a services approach, where a service can be thought of as a method of encapsulating business functionality behind a clearly defined interface that is technology agnostic and conforms to accepted practices.

4.4. IA-RM: Possible Solution Paradigms

There are several possible generic Information Architectures that provide semantic interoperability.
Without a firm foundation based on widely accepted standards a durable solution will be impossible. All solution paradigms can be based on open International standards.

Several solutions are presented:

- Data Dictionary paradigm;
- Message paradigm;

[20] HealthBase. Interoperability, Reference Architecture. Version1.0, December 2011

- Document paradigm;
- Two Level Model paradigm;
- Coding system paradigm / Ontologies

4.4.a. IA-RM possible solution: Data Dictionary paradigm

The Data Dictionary paradigm hinges on the need to map one field in a data base to an other field in an other data base. It is called a data base with meta-information about data bases, also.

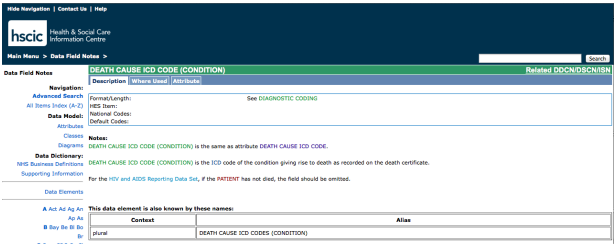
A data dictionary[21], or metadata repository, as defined in the IBM Dictionary of Computing, is a "centralized repository of information about data such as meaning, relationships to other data, origin, usage, and format."[1] The term may have one of several closely related meanings pertaining to databases and database management systems (DBMS):

- *a document describing a database or collection of databases*
- *an integral component of a DBMS that is required to determine its structure*
- *a piece of middleware that extends or supplants the native data dictionary of a DBMS.*
- *ISO/IEC 11179[22] (formally known as the ISO/IEC 11179 Metadata Registry (MDR) standard) is an international standard for representing metadata for an organization in a metadata registry.*

A Data Dictionary allows (as in any normal dictionary) to define lemma's where an item (many times called concept) is described in detail, its label as presented and some relationships are defined.
A Data Dictionary allows an implementer to find corresponding fields in an other data base. Or it allows customers to define terms in a dataset when they need to submit reports to a registry.

[21] http://en.wikipedia.org/wiki/Data_dictionary

[22] http://en.wikipedia.org/wiki/ISO/IEC_11179



An example is the NHS eHealth Data Dictionary.23

The details about the lemma (Death Cause ICD Code) are defined. Observe that although no codes have been specified it is possible to do so.

Observe that it is possible to create definitions using words, but computer processable output is impossible. And observe that it is impossible to define the structure of a dataset. In other words it is impossible to specify how many times an item is allowed to occur in a computer processable way.

Data Dictionaries can never represent a dataset fully (structure, constraints and the meaning of its components, including codes). Data Dictionaries can represent a defined dataset partially in the same way as a normal dictionary can define individual words. It is impossible to define in a computer processable way all possible relationships between the individual components.

Well known data dictionaries in health care are:

- NHS (England)
- NHS (Scotland)
- METeOR (NEHTA-AU)
- 3M

NHS (England)

The NHS Data Model and Dictionary for England[24] is the result of a long term interest in modeling healthcare.

23

http://www.datadictionary.nhs.uk/data_dictionary/data_field_notes/d/dea/death_cause_icd_code_(condition)_de.asp?shownav=1

24 http://www.datadictionary.nhs.uk/web_site_content/navigation/main_menu.asp

The NHS Data Model and Dictionary for England provides a reference point for assured information standards to support health care activities within the NHS in England. It has been developed for everyone who is actively involved in the collection of data and the management of information in the NHS.
This service consists of:

- Data model
  - Classes
  - Attributes
  - Diagrams
- Data Dictionary
  - Data Elements
  - NHS Business Definitions
  - Supporting information
- Clinical Content
- Data Collections

Since November 2013 the HSCIC of the NHS England deploys the EN ISO 13606 EHR communication standards next to a document management service based on ISO13119 Clinical Knowledge Resources - Metadata with extensions to that model.

NHS (Scotland)

The NHS in Scotland[25] maintains a Data Dictionary consisting of: The data Dictionary, Datasets and other standards (Social care, Accident & Emergency and Waiting Times).
All data is provided in a human readable format only.

METeOR[26]

METeOR is Australia's repository for national metadata standards for health, housing and community services statistics and information METeOR is Australia's repository for national metadata standards for the health, community services and housing assistance sectors. The system was developed by the Australian Institute of Health and Welfare to replace the previous repository, the Knowledgebase.

METeOR provides users with a suite of features and tools. These include online access to a wide range of nationally endorsed data definitions and tools for creating new definitions based on existing already-endorsed components. It has a strong focus on providing comprehensive user support and assistance.

[25] http://www.datadictionary.scot.nhs.uk/index.asp

[26] http://meteor.aihw.gov.au/content/index.phtml/itemId/181414

From a technical viewpoint METeOR operates as a metadata registry. This means METeOR is a system or application where metadata is stored, managed and disseminated. The registry aspects of METeOR have been based on the international standard for metadata registry - ISO/IEC 11179 - released in 2003.

Through METeOR you can find, view and download over 2,600 data standards. Using these standards will help you to:

- *avoid wasting resources creating similar standards*
- *base your information systems on nationally endorsed standards*
- *obtain data that is comparable across many different data collections*

METeOR provides powerful search facilities to help you find metadata quickly.

When the metadata you need does not exist, METeOR's metadata creation tools mean that creating quality metadata and getting them endorsed has never been easier.

All these services are available free of charge.

The METeOR system can be used to create Performance Indicators as National standards.

The development of indicators in METeOR improves quality, relevance, consistency and the availability of national information about the health and welfare of Australians. The drivers for standard development arise from the need for better information - whether it is statistical, administrative, clinical or other information. This ensures the data used in statistics is compatible and it facilitates National and International interoperability.

Information management groups and committees in the community services, health and housing assistance sectors have endorsed the use of METeOR for Council of Australian Governments (COAG) performance indicators to ensure compliance with the intergovernmental agreements of a National registry of Performance Indicators.

What is the value of METeOR to PI development?

- *Nationally maintained registry for PIs – the only one of its kind in Australia.*
- *Functional and user friendly templates that provide users with differing PI 'views'.*
- *Items in the registry are ISO-11179 compliant, therefore items can be and are used for national and international comparability e.g. cancer registries.*
- *Established governance processes.*
- *Related indicator information*
- *Conceptual frameworks*
- *Indicator sets and related indicators*

Indicator metadata item types

Indicators are endorsed as standards by the relevant registration authority. There are a five metadata types that make up the suite of indicator templates.

- *Indicator set*

- *Outcome area*
- *Indicator*
- *Quality statement*
- *Data source*

The METeOR Data Dictionary allows the searching of defined concepts, datasets, associated code sets and exports to word and PDF formats.

The 3M Healthcare Data Dictionary [27] (HDD) is a controlled medical vocabulary server that has been continuously expanded and maintained for over 15 years. The 3M HDD makes it possible to map and manage medical terminologies, integrate content and standardize healthcare data. The technology allows organizations to transmit and receive accurate, actionable patient data across systems and applications, regardless of where data originates.

3M HDD is an integration tool that primarily focusses on interoperability of individual concepts and their codings. It is an 'active' Data Dictionary because it allows the run-time translation of codes when systems exchange data.

4.4.b. IA-RM possible solution: Message paradigm

Message passing[28] is a technique for invoking behavior (i.e., running a program) on a computer. In contrast to the traditional technique of calling a program by name message passing uses an object model to distinguish the general function from the specific implementations. The invoking program sends a message and relies on the object to select and execute the appropriate code.

The message paradigm is a specification that is designed to update data base fields of the receiving party with data sent by the sender. In healthcare examples of standards are: UN/Edifact messages based on European CEN message standard are deployed in Europe and HL7 v229 and v3 in the USA and elsewhere.
Edifact messages in healthcare are mostly based on CEN message standards they are widely used in Denmark, and the Netherlands.
HL7v2 are widely used in healthcare institutions.
Initially the exchange format of messages is position dependent. The place in the message provides the semantics.
HL7v3 messages more explicitly model the semantics because they use the Reference Information Model on which messages are based provide semantics.

---

[27] http://solutions.3m.com/wps/portal/3M/en_US/Health-Information-Systems/HIS/Products-and-Services/Products-List-A-Z/Healthcare-Data-Dictionary/

[28] http://en.wikipedia.org/wiki/Message_passing

[29] For examples see: http://www.ringholm.de/docs/04300_en.htm

A distinct feature of the messaging paradigm is the fact that the message standard encompasses several viewpoints as defined in the ISO RM/ODP standard: Enterprise, Informational, Computational and even the Engineering viewpoints.
In addition each of these methods allow too many degrees of freedom to model the same thing.
Each implementation of a message in an IT-system is a unique implementation.
All together this results in solutions that need a lot of engineering resources to create interoperability at all these RM/ODP layers before the message standard can be deployed in regions or countries. In fact an other organisation is needed to profile these standards and test the vendor engineered solutions. Integrating the Healthcare Enterprise (IHE30) develops these profiles and organises Connectathons.
Once in a region or country these profiles are implemented and tested in vendor systems any change needs to traverse the IHE process again.

The consultancy firm Gartner over the years has analysed this paradigm and have declared this method not viable in the long run because of its problems. HL7v2 and Edifact will be used for considerable time for integration inside IT-systems. HL7v3 received the deprecated status because of the implementation problems as the result of inherent problems of the deployed Reference Information Model, modeling methods, the too many degrees of freedom that is allowed and the extensive problems and resources needed for the implementation.
Since CDA is predicated on HL7v3 this development will influence the future of HL7 CDA.

A recent development is FHIR31. Fast Health Interoperable Resources (FHIR) is on its way to become an HL7 (ANSI) standard. It is presented as:
'… a next generation standards framework created by HL7. FHIR combines the best features of HL7's Version 2, Version 3 and CDA® product lines while leveraging the latest web standards and applying a tight focus on implementability.'
FHIR is work in progress.

4.4.c. IA-RM possible solution: Document paradigm

The document paradigm developed the notion to use the message paradigm to produce documents that contain data from an EHR and send it to an other EHR-system. HL7 CDA is such an example. Documents can be represented that are based on the HL7v3 Reference Information Model. The HL7v3 CDA document paradigm suffers from all negative aspects that the (HL7v3) message paradigm has.
Gartner[32] predicted that the document CDA paradigm has proved itself and will be deployed for considerable time. While ate the same time deprecated HL7v3 technology as 'obsolete before plateau'.

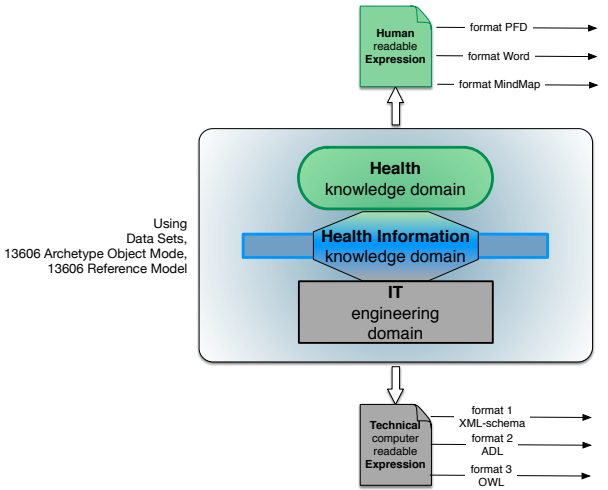
[30] http://www. IHE.net

[31] http://www.hl7.org/implement/standards/fhir/summary.html

[32] Gartner Hypecycle published 2009

HL7 CDA is selected by the European Member States as one of the European Interoperability Framework base standards.

4.4.d. IA-RM possible solution: Two Level Model paradigm



This Two Level Modeling paradigm is the method behind the EHR exchange standard. This EN ISO 13606 EHR-communication standard is developed by CEN/ISO. In this paradigm at least two models are needed to define artefacts. One model deals with all aspects of the structure of any document plus archiving facilities. A second model allows the production of artefacts as constraints on the first model. Any artefact will express what needs to be documented and exchanged according to user groups. Any artefact will be conformant to the first (Reference Model) and second model (Archetype Object Model). When the artefacts are produced based on a third model that defines how artefacts are structured the artefacts are very expressive and modeled with few degrees of freedom.

A substantial advantage of this paradigm is the fact that because of the precise models that guide the production of semantic interoperability artefacts these artefacts on one hand can be used by healthcare providers and at the same time can be used by IT-systems. This reduces the amount of resources needed to implement this paradigm and at the same time the model driven nature allows very fast change at local, regional and national levels.

The EN ISO 13606 EHR Communication standard fulfills the requirements of the ISO standard[33] on requirements for EHR architectures and allows the expression of the Patient Mandate (ACL) that defines who has access to any data point.

In a recent development this EN ISO 13606 EHR Communication standard is aligned with two other important standards in a process named Concurrent Use: CEN/ISO 13940: System of Concepts for Continuity of Care and CEN/ISO 12967.
The joint use of these three harmonised standards results in specifications that deal with the Enterprise, Information and Computation RM/ODP viewpoints and cover the domains of health care processes, service interfaces inside an EHR-system and the EHR data that is used.

4.4.e. IA-RM possible solution: Coding / ontology system paradigm

This paradigm hinges on the exclusive use of an ontology and codes such as SNOMED-CT for each datum in a message or data base. For this purpose SNOMED must be able to define codes for data in their complete context. IHTSDO has worked on a part of SNOMED where context information can be combined with data. E.g. It allows to code for Systolic Blood pressure of the patient and a different code for the same measurement performed on a relative.
In the case of datasets that need a precise specification of all aspects of the context it must be possible to find all the needed SNOMED codes. In the course of the production of the epSOS dataset it became clear that SNOMED is lacking concepts and codes.
In addition a precise specification of datasets can not be done without the possibilities to create a structure and define precise constraints on allowed numbers, letters, texts, etc.

In this context the NHS Logical Record Architecture[34] (LRA) must be mentioned. This solution uses a subset of EN ISO 13606 and makes extensively use of SNOMED-CT. And will encounter the same problems that not the complete context (epistemology) of data can be specified.

Recently the NHS-HSCIC changed their approach and now follow fully the EN ISO 13606 standard.

4.4.f. Summary: Comparison of the paradigms

All presented solution paradigms have been analysed. The results are presented in the table below.

33 ISO 18308: Requirements for an Electronic Health Record Architecture

34 http://www.uktcregistration.nss.cfh.nhs.uk/trud3/user/guest/group/0/pack/12

Paradigm	Scope	Degrees of freedom	Standards Conformance 18308, 13940, 12967 22220	Patient Mandate at the data item level	RM/ODP viewpoints	Data Expressivity	Human and computer processable artefacts	Resources Needed	Flexibility	EIF Base Standard
Data Dictionary	Define and manage data elements	+++	111179	-	Enterprise Information	-	+	+	-	-
Messages (1990-)	data base fields updating	+++	-	-	Enterprise, Information, Computation, Engineering	+++	-	+++	-	-
Documents (1996-)	document exchange between IT-systems	++	-	-	Enterprise, Information, Computation, Engineering	++	-	+++	-	+
Two Level Model (2001-)	EHR communication	-	+	+	Enterprise, Information, Computation,	++++	+	+	+	+
Coding Ontology	reference terminology and supporting ontology	± Managed by an ontology	-	-	Enterprise Information	-	±	+	±	-

The Two Level Modeling paradigm provides the best fit to the requirements for an IA-RM.

Require -ment number	Requirement text	Comment	Data-Dictio-nary	Messages	Docume nts	Two Level Model	Coding / Ontolog y	Remarks
1	The solution must be based on open International standards	The solution needs to be effective, efficient, be usable in the European market context of procurements and opportunities in other countries	+	+	+	+	+	
2	The solution must be able to facilitate the deployment of National and Regional policies	In Europe European standards play a formal role in the deployment of National policies for legal and ethical matters	±	+	+	+	±	Data dictionary and Codings only express part of what is needed
3	The solution must be usable in procurement processes	European standards play a formal role in the creation of a common market	+	+	+	+	+	
4	The solution must allow an evolutionary process from present systems to new systems in the future	Products and services have an economic life cycle. Occurred investments need to be recognized	+	+	+	+	+	
5	The solution must allow existing functional exchange/interface formats to be supported	Groups of healthcare providers that exchange data using existing message based solutions will have a growing need to deal with the new National Exchange formats.	±	+	+	+	±	Data Dictionary and Ontology can not express all of that what is needed (see R3)
6	The solution will consist of a set of common and sharable services	The Irish ICT National Integrated Services Framework must have services can be used by all stakeholders	+	+	+	+	+	
7	The solution must be durable and affordable	Implementation experiences learn that generic interfaces that support both HISA as well as EHRCom standard are not more expensive to develop as proprietary one-to-one interfaces, but generate a considerable cost-efficiency by their re-usable nature and flexible deployment.	+	±	±	+	+	Messages and Documents need substantial resources (Time and €) to produce, implement and maintain
8	Components of the solution must be governed	Common and shared services that express components of agreements for exchange between stakeholders need to be trusted and be durable	+	+	+	±	+	Two Level Model is a model driven solution Technical conformance is automatically done by parsers at design time.
9	The solution must be able to define the components of the 'Semantic Stack' (datasets, codes from codings systems) as a common service.	Published agreements via a Data Dictionary and that include the agreed codes from coding systems fully normalize the data exchange between stakeholders	±	+	+	+	±	Data Dictionary and Codings only express part of what is needed

Require -ment number	Requirement text	Comment	Data-Dictio-nary	Messages	Docume nts	Two Level Model	Coding / Ontolog y	Remarks
10	The solution must support the creation and maintenance of a Data Dictionary as a common and shared managed resource that helps unify datasets and the unified expression of these datasets	Published agreements via a Data Dictionary must be made available as a service users/ stakeholders can interface with.	+	-	-	+	+	Messages and Documents need Implementation guides and Profile documents. They do not function as Data Dictionary.
11	The solution must support the fact that EHR-data must be exchangeable fully or in part	Parts of the Patient Record must be exchangeable between stakeholders, but also complete records.	+	-	-	+	-	Messages, Documents have a scope smaller than the complete EHR.
12	The solution must allow groups of stakeholders to co-operate in the production of their dataset and codes from coding systems		+	+	+	+	+	
13	The solution must be able to respond in a short time to new requirements	Solutions that need many months or years to get from expression of new requirements to the implementation stifle healthcare innovation and healthcare reforms	+	-	-	+	+	Messages and Documents are resource intensive.
14	The solution must support the use of clinical pathways, and protocols, for cooperating stakeholders	EHR data in IT-systems and in exchanges between stakeholders document the provision of healthcare. Healthcare is a complex interplay of many processes at various levels. The data that is documenting these processes and is reporting about it, must reflect the fact that the data is part of these processes.	+	±	±	+	+	Two Level is aligned with CEN/ ISO System of Concepts. Others are not. Messages and Documents support processes to a degree, but are not aligned with ContSys
15	The common services will produce artefacts that at the same time must be readable and understandable by healthcare providers and be readable and usable by the IT-industry	One expression for the artefacts in the common and shared services that can be used in both the health and technical domain secure consistency and easy and fast implementation	H: +	H:+ T:+	H:+ T:+	H:+++ T:+++	H:+ T:+++	Two Level Model is capable of automatically generating: tree's, mindmaps, HTML screens, Excel, Schematrno, XML data instantiation, ADL

A simplified evaluation scheme is used in the Overview documents.

Based on the statements made in the mini-tender introduction text, objectives presented, the need for a proof of concept and the outline of the requirements the following decision tree is used to structure the work in LOT 4.

The next summary tables are the result workshops, discussions and questionnaire.

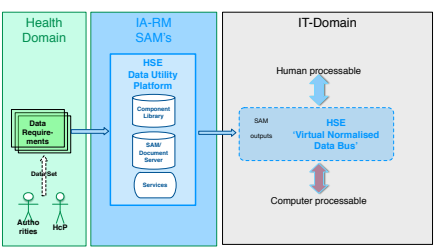
	Open Standard	RM/ODP Viewpoint	Scope EHR	Implementation Maintenance cost
HL7 v2 messages	+	-	-	-
HL7 v3 messages	+	-	-	-
HL7 v3 CDA documents	+	-	±	-
EN ISO 13606 EHR Communication	+	+	+	+

The table provides the rationale leading to the selection of the EN ISO 13606 as the standard for the HSE ISF Information Architecture Reference Model.

Criteria	Justification
Open International standards	All listed possible standards are open international standards. Alternatives that were not listed are: **HL7 FHIR** fails the criterium because it is not a formal CEN and ISO standard. **OpenEHR** fails the criterium because it is a proprietary specification that shares the Two-Level-Modelling approach with the EN ISO 13606 (with a different Reference Model, Archetype Object Model and Data Types). OpenEHR is not a formal open international standard.
Scope is to create a library of re-usable structures (SAM's) that fit the RM/ ODP Information Viewpoint only	The RM/ODP standard defines 5 Viewpoints: Enterprise, Information, Computation, Engineering and Technology. Message standards and CDA do not restrict themselves to the Information Viewpoint. They extend into the Engineering Viewpoint. The EN ISO 13606 EHR communication standard restricts its scope to the Information Viewpoint. Archetypes and Templates produced express the datasets, data requirements in a human and computer processable way.

Criteria	Justification
Scope is EHR communication	Message standards like HL7 v2, HL7 v3 and the HL7 v3 CDA document standard fail this criterium since they do not have a full scope to transport data between EHR-systems. The HL7 message standards have a restricted scope of exchange between components of hospital IT-systems .The scope of HLv3 CDA is wider but is not covering the complete EHR, since the document structure of HL7 v3 CDA is a subset of the structure as defined by the EN ISO 13606-standard.
Implementation and Maintenance cost	HL7 v2 and v3 messages and HL7 v3 CDA documents share the same necessary implementation process when they are deployed. First IHE profiles of these standards are produced such that they represent the local use cases. Hereafter in the second phase all vendors that claim compliance need to be tested in IHE Connectathons. Considerable resources (time and funds) are needed for both steps. Any change in the standard or the profiles necessitate to follow all steps again. EN ISO 13606 EHR communication standard allows an agile, flexible, process because of the Two Level Modelling Paradigm that is used. A Model Driven local deployment. using predefined re-usable components (archetypes/templates), is possible. specified by the CEN/ISO EN1360

4.5. IA-RM: Description



The figure shows the high level picture of the proposed IA-RM architecture.

There are three domains:

- **Health care domain**: Data requirements collection resulting in agreements between communicating partners that describe what data will be shared.

- **IA-RM domain**: Subject Area Models (SAM's) that define in a human and computer readable format a transcription of the data needs in a standardised way.
- **IT-Domain** (Procurement / Implementation / deployment) Defined SAM's and resulting outputs are used in procurement by HSE. In addition the IT-vendors use the SAM's and resulting outputs to implement the data requirements in their IT-systems and conform to the procurement requirements.

The Information Architecture Reference Model will be a library of common, shared, re-usable, Subject Area Models (SAM's) and a tooling environment that will act as governed resource.

Subject Area Models (SAM's) and how they are defined will be described in detail in an other chapter of this document.
The 'Data Dictionary' functionality will be described in an other part of this Deliverable.

Stakeholders define their data requirements as 'datasets' that are transformed into Subject Area Models (SAM's) in the HSE Toolbox Environment.
The HSE Toolbox Environment is a set of applications/services:

- Content Model Library that holds re-usable components of Content Models as Subject Area Models.
- Content Model Library holds the Subject Area Models, also.
- Data Dictionary an additional service to help maintain consistency between the Subject Area Models. It defines at any level of detail the definitions for the data points defined in the Subject Area Models including bindings to coding systems.
- The HSE Toolbox allows groups of users to cooperate on the production, validation of the Subject Area Models.

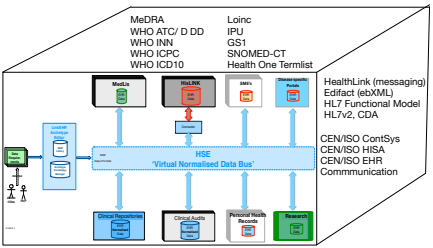
The HSE Toolbox can produce as validated and governed output various methods to display the Subject Area Models as re-usable Semantic Interoperability Artefacts in versions for use by health

Important aspects for the successful deployment are:

- Explicitly no message standards will be prescribed.
- Explicitly no requirements will be imposed on existing solutions to make changes to their established exchange formats and methods for internal storage, retrieval and processing in their IT-systems.
- The single requirement for healthcare providers and organisations will be that their (procured) IT-systems conform for exchange to the described information viewpoint (using the IA-RM and subsequent SAM's and outputs) when communicating outside their present jurisdiction.

The IA-RM and its components will create a flexible level playing field for all existing and newcomers to the market when they need to exchange data with the Governmental and its Agencies.

4.6. IA-RM and stakeholders



The IA-RM defines the HSE Toolkit that produces Semantic Interoperability Artefacts (Outputs) that the various stakeholders can use to express clinical datasets as Subject Content Models. It creates a 'green field' for all actors; an eHealth ecosystem where all actors Irish healthcare providers, their IT-systems and IT-industry can play their roles.

Existing systems many times use other paradigms for their communications, such as HL7v2, and HL7v3 CDA R1 and R2 formats in conjunction with IHE profiles. IT-systems use different methods to store and retrieve data in the systems.

These existing systems and solutions will NOT have to disband these existing methods for exchange. In order to comply and make use of the SAM's and outputs for exchange they must install an additional interface that is capable to present data in normalised format to those outside their jurisdiction.

Or alternatively these existing IT-systems provide the data in the existing formats for re-use. The receiving sites then need an Integration Engine that does the transformation to the normalised data format.

The proposed IA-RM makes an evolutionary change process possible when deployed as a desired way forward. This IA-RM, with the SAM and outputs plus the Data Dictionary is inline with the wishes as expressed by the stakeholders in the Work Shop discussions and questionnaire.

4.6.a. IA-RM and Healthcare providers

Co-operating healthcare providers can produce and discuss their data requirements, also called (Clinical) Datasets using the HSE Toolbox. They can use all kinds of methods (text, mind maps, excel spreadsheets, …) they prefer to define their datasets. Supported by Health-IT specialists, that most often are trained nurses that translate datasets into Content Models. SAM's are produced, discussed and validated. After validation the SAM's will be published.

SAM's specify the data points but also their context and associated codes from Reference Terminologies plus Classifications and own local code lists. In this way SAM's are a resource where local codes can be mapped to Reference Terminologies and Classifications.

The open standard based, model driven architecture of the HSE Toolbox makes it possible that new requirements can be implemented in a short period of time.
The HSE Toolbox Environment holds the data requirements that can be used in procurement of systems. It indicates what, next to functional requirements, data points are expected that the next IT-system must be able to handle.
Integration Engines make it possible that these IT-systems can read and write data according to the normalised format used in the 'Normalised Data Bus' and existing investments in IT-systems can be leveraged. The alternative is the implementation of several new messages supported by those IT-systems.
The above makes an evolutionary change process possible as a desired way forward to deploy the IA-RM and Data Dictionary outcomes as stated by stakeholders in the Work Shop discussions, which statements were overwhelmingly supported by the questionnaire outcomes.

4.6.b. IA-RM and Organisations

Several organisations of stakeholders exist such as: hospitals, communities that use and exploit a communication platform, organisations that provide disease specific platforms and registries, and organisations that do research.
All of these organisations that need to use, re-use and exchange health data can use the HSE Tools to produce artefacts that express their data needs that can be used in the procurement of new IT-systems.
Integration Engines make it possible that these existing IT-systems can read and write data according to the normalised format in the 'Normalised Data Bus'.

4.6.c. IA-RM and Authorities

Authorities are one of the stakeholders that could use the possibilities that the HSE Toolbox and the 'Normalised Data Bus' will provide. When legislation or policies demand the reporting of prescribed datasets, then the proposed solution will allow them to define these datasets and how the data needs to be reported. The 'Normalised Data Bus' will enable the stakeholders to comply with the demands.

Semantic Interoperability artefacts will enhance the effectiveness of the National legal and policy requirements. These published artefacts are: datasets, SAM's, outputs are made available for use in procurement.

The demand to make data available in the prescribed data format via SAM's and outputs plus supporting services, together with the use in procurement will be a minimal but sufficient cost-effective method to create an eHealth Infostructure.

No substantial investments are necessary to produce, publish, test and maintain message standards in all implementations with healthcare providers.

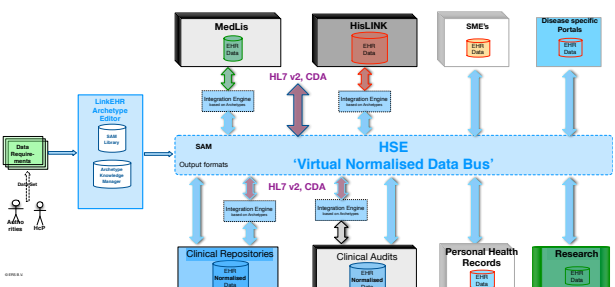
4.6.d. IA-RM and Research

Re-use of data for research needs access to data stored in EHR-systems. Data that is optimised for supporting the provision of care to patients. IT-systems that healthcare providers use store and retrieve data using vendor specific proprietary methods. When data is re-used many times message standards such as HL7v2 or HL7v3 CDA formats are used to provide the data.
Because research questions vary, this implies the need for interfaces that create flexible access to data in IT-systems. In general message standards are inflexible and resource intensive.
A 'Normalised Data Bus' allows the flexible definition of the research payload such that healthcare providers and their organisations can participate more easily in research projects.

4.6.e. IA-RM and IT-Systems

IT-systems connected to a 'Normalised Data Bus' operate in a plain level playing field for existing industries but also for SME's, when the Data Bus is based on open International standards. Subject Area Models (SAM's) will express in detail the data needs of health actors. The corresponding SAM's and the output formats generated from it can serve in procurements and they can be used for implementing these requirements in the systems interface or inside IT-systems. The IT-domain will be optimally informed and facilitated.

4.7. IA-RM: Summary



- The Information Architecture Reference Model is a high level, abstract, logical, model that serves as the point of departure -the basis- for developments in a region, in this case Ireland.
- The IA-RM guides the development of the Irish INFOstructure that supports all actors in healthcare when they want or need to exchange data about patients and the healthcare processes.
- The IA-RM creates a 'green field' eHealth ecosystem for all actors.
- The IA-RM has chosen to focus on the Information Viewpoint meaning on the specification of datasets, called Subject Area Models, that represent the data needs for healthcare actors.

- SAM's specify the data that will be exchanged, but increasingly, the data that will/can be stored and retrieved inside IT-systems.
- The technical output formats of the IA-RM based SAM's will support flexible and cost effective implementation in IT-systems.
- SAM's will be produced in a central resource as governed environment.
- Output from that resource are called Semantic Interoperability Artefacts (Outputs) and represent SAM content in various human en computer processable formats.
- The IA-RM, SAM's and outputs allow an 'HSE Normalised data Bus'.
- An evolutionary process is fostered where SAM's as part of procurement and remuneration processes will guide future developments of re-use of data for joint integrated care, reporting, audit registries, and research.
- The IA-RM creates a level playing field for small and large healthcare providers and organisations, plus small and large IT-vendors.

5. Subject Area Model

This chapter corresponds to part 1.2 of the project Deliverables.
Using the epSOS patient summary dataset a Subject Area Model (SAM) was produced using a selected technology. The SAM was demonstrated in various presentations.

5.1. SAM: Introduction

A Subject Area Model (SAM) is a specification that reflects as the result of consensus the data requirements of a group of Healthcare Providers in their context and for a specific purpose.

5.2. SAM: Why are SAM's important

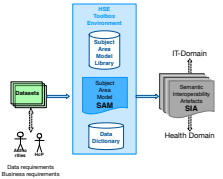
Other names for SAM are Dataset or Data standard or Detailed Clinical Model or Semantic Interoperability Artefact.

Why use Subject Area Models (data standards)?[35]

- *Making data count*
- *Data standards promote the quality accuracy interpretability reliability relevance interchangeability transparency currency accessibility coherence and comparability of data and information.*
- *Without data standards there is the potential for data to be of poor quality. Data may fail to measure what it is supposed to measure or not be comparable across collections or over time. Decisions based on poor quality data affect us all whether it is hospital services or delivery of services in the community.*
- *Data standards enable consistent and comparable reporting of information about services and people including describing what services are available where services are located to whom they are delivered by whom are they delivered and when how much they cost and what happened as a result of delivering services.*
- *By making endorsed data standards readily available users are assured that they can use these standards with confidence and that they will enable the maximum re-use of their data for future research 'create once use often'.*

[35] Text adopted from 'National data dictionaries', National Community Services Data Dictionary , version 5, Australia

Subject Area Models (SAM's) are built according to data standards, and using other open International Standards.
SAM's are the translations of datasets defined by healthcare providers into a computer processable format. The SAM's have to be produced, validated by their stakeholder communities. A governing organisation needs to control the HSE Toolbox as managed common shared public resource.



SAM's define the data points that stakeholders need and the associated codes from coding systems. The SAM's, when exported, are called Semantic Interoperability Artefacts that can be presented in a healthcare friendly way. But also as a technical expression that can be used by IT-vendors and other IT-specialists.

In order to keep a consistency between the SAM's they are built from predefined shared building block from the library. This arrangement secures that all SAM's define the same things in the same consistent way. E.g. The demographics, date of births, what is documented about medication, lab tests, diagnosis, findings, etc.
The re-use of common building blocks helps to create uniform data in the 'HSE Normalised Data Bus'. In addition a Data Dictionary helps secure an overall consistency of the SAM's and resulting outputs.

Individual healthcare providers, groups of providers, their organisations and governmental organisations have a need to re-use and share data about patients and their treatment. Data that is stored in their IT-systems.

Each vendor is using proprietary methods to deal with the data inside systems. Each user group has in s specific context a need for specific data. The context can be for instance:

- a digital prescription sent to the pharmacy
- an order for lab tests and the reporting of the results
- a referral or discharge letter
- obligatory of voluntary reporting to registries or for auditing purposes
- management reports
- research internal or external
- etc.
- When an IA-RM and supporting services are available they can be used to collect the data requirements a Subject Area Model (SAM) also known as Dataset. Many times this dataset is described in text and tables.
  Using the IA-RM supporting services the verbal specification is converted into human and machine readable formats. After a validation phase the consensus is made available to the public.

5.2.a. Healthcare provider and their organisations

An individual Healthcare provider or group of providers can use the published SAM as a specification when selecting a new IT-system or demand updates for their present system.
Healthcare providers can use obligatory or voluntary SAM for auditing or management reporting to influence their vendor.

Agreed general or ad-hoc SAM's when implemented can allow the provider to participate in research.

5.2.b. Health agencies

Governmental health agencies have a need for data from Patient EHR's for audit and policy reasons. SAM's will define their data needs that when published inform Healthcare Providers and their organisations to implement these requests for data.

5.3. SAM: International Developments

Several developments are ongoing with respect to the creation of Subject Area Models that describe the data needs of groups of healthcare providers.

The most important developments are:

- Detailed Clinical Models
- General Purpose Information Components
- NHS-HSCIC
- NEHTA (AU)
- epSOS Patient Summary specification / EU-Recommendation
- Clinical Information Modeling Initiative

5.3.a. DCM

A detailed clinical model is a standalone information model designed to express a clinical concept in a standardized and reusable manner. It documents clinical information as a discrete set of precise clinical knowledge for that concept. This is an ISO 13972 specification for defining detailed clinical models entitled "Health informatics - Detailed Clinical Models":

- 'Part 1: Quality processes regarding detailed clinical model development, governance, publishing and maintenance'. Committee draft in 2011.
- 'Part 2: Quality attributes of detailed clinical models'

The ISO 13972 DCM specification describes the structure, its characteristics and production process. The output of that process is a text describing in detail the data needs.

Only groups of health experts can validate these structured documents. The DCM specification discusses the governance process.

For patient safe semantic interoperability it is essential that all DCM's are designed and produced in a standardised way as specified.

Only in HL7 work is underway to define DCM's.

5.3.b. GPICS

General Purpose Information Components are defined in the CEN EN 14822:2005 4 part standard. It is derived from the data needs of 15 years of European message standardisation in healthcare. The GPICS cover the clinical and non-clinical domains and are expressed as HL7 v3 derived artefacts.

5.3.c. NHS-HSCIC

The NHS for many years worked on their SAM variant using the Logical Record Architecture as that is a profile of the EN ISO 13606. Since November 2013 they build SAM's using the full 13606 Reference Model.

5.3.d. NEHTA[36]

Using the proprietary specification from 'openEHR' NEHTA is building a library of openEHR archetype artefacts.
More openEHR based collections can be found at:
http://www.openehr.org/wiki/display/healthmod/Clinical+Knowledge+Manager

5.3.e. epSOS / EU Communication

epSOS is a European project run by Member States. It defined and implemented a specification for cross border exchange of the Patient Summary and medication/e-Prescription. The project is about to be terminated.
The Patient Summary SAM has been defined.

The most tangible output is the EU Communication[37]:
Guidelines on minimum/non-exhaustive Patient Summary dataset for electronic exchange in accordance with the Cross Border Directive 2011/24/EU as published on 19 November 2013.

5.3.f. CIMI

The Clinical Information Modeling Initiative exists for two years. Important members are: US: DoD, VHA, Mayo Clinics, InterMountain, …. . Other actors are: NHS-England, Singapore, NEHTA, CEN/TC251, EN13606 Association, … .

5.3.g. Summary

In summary several competing Subject Area Model sources based on various standards and non-standards and various interpretations of the (Non-) standards are available. There is no authoritative source and solution at this moment.

---

[36] http://www.openehr.org/wiki/display/healthmod/NEHTA's+Clinical+Knowledge+Manager

[37] guidelines on minimum/non- exhaustive patient summary ...

It is a safe assumption that solutions that are open International standards based will be more successful than others.
It is a safe assumption that those groups that have more important actors than others will be more successful.
It is a safe assumption that whatever the final international consensus all of the mentioned can and will provide ideas and input as starting points for collections of Subject Area models in the (near) future.

5.4. SAM: Description

The SAM is the result of a consensus process and informed by the needs of healthcare actors.
The specification that describes the SAM in detail has two content related aspects:

- Health aspect and
- Technical aspect

Each of these aspects will be discussed.

5.4.a. SAM: Health content aspect

The collection of SAM outputs contains formats that can be processed by humans.

DCM

This aspect is known under the name Detailed Clinical Model (DCM). A detailed clinical model is a standalone information model designed to express a clinical concept in a standardized and reusable manner. It documents clinical information as a discrete set of precise clinical knowledge for that concept.

The ISO 1397238 DCM specification describes the structure, its characteristics and production process. The output of that process is a text describing in detail the data needs.

Only groups of health experts can validate these structured documents. The DCM specification discusses the governance process.

For patient safe semantic interoperability it is essential that all DCM's are designed and produced in a standardised way as specified.

SAM

In the next step the DCM will have to be converted in a SAM. An health informatics experts with knowledge about DCM's, Archetypes, Coding systems and the described service processes will convert the dataset (DCM) into one or more EN ISO 13606 archetypes.

Groups of relevant experts will receive the SAM as Excel spreadsheet, as MindMap, or mock-up screen to inspect the translation of the DCM into the SAM.

38 ISO/DST 13972 Health informatics -- Detailed clinical models, characteristics and processes

For patient safe semantic interoperability it is essential that all SAM's are designed and produced in a standardised way as specified.

Increasingly important will be the fact that SAM's allow the documentation of a clinical fact in its complete context. SAM's could be constructued using one standardised pattern known as the Semantic Interoperability Artefact Modeling Method[39] (SIAMM). SIAMM reduces the number of degrees of freedom and prescribes that each clinical fact can be documented including its full semantic rich context.

SAM's are constructed (re-)using predefined components, thereby guarantying that the same concepts are modeled in a uniform way. The general SIAMM pattern is annotated with relevant SNOMED and LOINC codes.

5.4.b. SAM: Technical content aspect

The collection of Semantic Interoperability Artefacts (Outputs) contain a series of artefacts that software engineers can use immediately in the process of creating messages for reporting or implementing it directly into their IT-system or use it to prepare for third party validation/ certification.

Actors that validate, qualify or certify claims by IT-system vendors that they are conformant to specific SAM's need those technical artefacts such as: ADL1.4 Archetype and Template formats, XML instantiations of data, XML-Schematron formats.

5.5. SAM: Summary

Overseeing many international developments and taking in consideration the results of the chapter on the Information Architecture Reference Model the next conclusions can be drawn:

- Subject Area Models (SAM's) using EN ISO 13606 archetypes describe all the detailed data needs of (groups of) Healthcare actors that need to store and exchange data.
- SAM's are produced as the result of consensus.
- SAM's are built using a series of supporting standards and codes from relevant coding systems.
- SAM's are produced using a centrally governed resource consisting of an editor, a document manager, terminology servers/services, and acting as a collaboration tool.
- SAM's as usable output produce Semantic Interoperability Artefacts in various formats that humans and computers can process.
- SAM's need to be governed and validated before use.

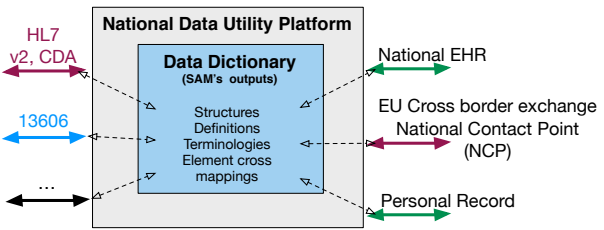
[39] Developed by ERS B.V. and will be submitted for inclusion in the EN/ISO 13606 standard

6. Determination of potential participating technical systems

6.1. Technical Systems: Introduction

This chapter corresponds to the Project Deliverable 1.3.
Stakeholder groups were questioned using questionnaires and workshops to validate the proposed approach and possibilities for integration. See the annex.

6.2. Technical Systems: What is the importance



In general it is important that the approach taken and the results presented are validated by those actors that will be influenced by the decision to create an IA-RM and SAM's per knowledge domain.

The EN ISO 13606 standard is designed to accommodate the transfer of complex patient data form any EHR to any EHR including Registries. Tools exist that can transform proprietary data into any other format. The modelling paradigm (Two-Level-Modelling) that is used in the standard enables this versatile deployment. Model 1: the EN ISO 13606 Reference Model and model 2: the Archetype Model.

Proprietary EHR-systems that make use of other standards for communication can view, inspect, comment the SAM's and can use the technical output formats in the deployment of the HSE ISF SAM's and generated artefacts.
Human readable outputs at this moment are: Tree views, Graphical MindMaps, Excel spreadsheets (that can be input for any data base), generated mock-up screens.
Technical outputs at this moment are: ADL 1.4 files, XML files, Schematron, XML data instantiations.

The SAM technical outputs can play a role in a Connector / Interface that allows system with proprietary data formats to exchange data. Based on experiences by HSE and ERS it can be concluded that implementation costs occurred by vendors for a generic Connector / Interface based on EN ISO 13606 are the same as for one HL7v2.x message, only. Operational cost are reduced because of the possibility for a, effective and efficient Model Driven implementation of the Connector / Interface.

One of the specific connectors is for the cross-border exchange of the epSOS datasets. For this purpose a National Contact Point is necessary to take part in European developments.

6.3. Technical Systems: Summary

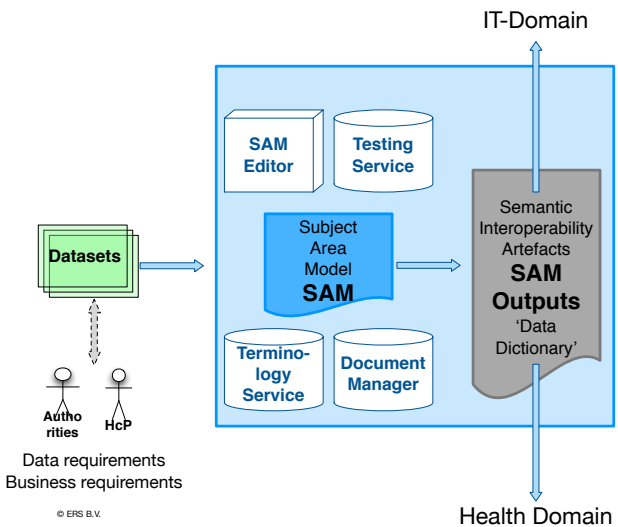
The Annex chapter and Workshop reports indicate a solid support for the approach to normalise the exchange of data between health actors and the fact that an evolutionary trajectory is possible to and from existing exchange formats.

7. Tools supporting the Subject Area Model

This chapter collects the various tooling aspects in the Project Deliverables.

7.1. Tool: Introduction

This chapter will describe the HSE Tooling Environment (HSE Data Utility Platform) and the artefacts that are produced.



The Subject Area Model describing the data needs / dataset for a particular use case is produced using a SAM-Editor.
Many times this will be (what is called) a Template). Templates (SAM's) are constructed using predefined standardised building blocks called archetypes.
All archetypes and templates need to use as many as possible existing codes from existing coding systems. Terminology Servers allow the finding and use of all these codes inside archetypes/ templates.
All archetypes and templates can be used to general various output formats such as: ADL, MindMaps, Excel spreadsheets, XML-Schematrons, XML data instantiations, implementation guide, etc.
All these artefacts need to be managed in the Document Manager in order to secure consistency between them.
All artefacts (SAM's, templates/archetypes and supporting materials) need to be governed and validated. Technical validation aids in the Testing environment will support the governance of the artefacts that are published. In addition the Testing Service will allow IT-software developers to test their implementations against and support the technical deployment of the SAM's.
Finally the Tools and all artefacts produced need to be made available in a Collaboration environment, so all actors involved can work on all artefacts.

7.2. Tool: Why are they important

The Two Level Modeling paradigm that is model based supports the production of a SAM Editor to define in high detail and in a very structured way Subject Area Models. Without standards normalised human and computer processable output is impossible.
SAM's rely on many codes from coding systems. One localised environment where the codes and codings systems are made available is indispensable.
The re-usable building blocks to construct SAM's need a document manager tool. The collection of SAM's and outputs in the various output formats need a document manager tool, also. Without such a document manager tool the complex collection of documents can not be governed.
All Healthcare actors that produce and make use of the SAM and outputs need an environment, as service, where they can search for documents they need to discuss, validate or use or make comments about. Around the document manager a controlled collaboration environment needs to be in place.
Without the described governed set of tools and services a functional centrally governed INFOstructure supporting semantic interoperability is impossible.
Tools are indispensable for the creation, maintenance, publication and support of the 'Normalised Data Platform'.

7.3. Tool: SAM Artefact Editor

The SAM Artefact Editor is a tool that builds the specification for any Subject Area Model.

Requirements for the SAM Artefact Editor that have to be fulfilled are:

1. *standards based;*

2. *must produce artefacts that can be understood by health domain experts and IT domain experts;*
3. *must produce artefacts that must be computer processable by healthcare domain experts and IT-domain experts;*
4. *allow the precise expression of any data structure;*
5. *allow to attach one or more codes to nodes in the data structure;*
6. *allow to attach to leaf nodes: components like: codes, multimedia, ordinal scales;*
7. *allow to express in leaf nodes constraints on possible data values, data types, and code values;*
8. *allow the expression of any clinical statement (quantitatively, semi-quantitatively and qualitatively) plus its complete context.*

On top of these tooling requirements additional requirements pertain to the artefacts produced. These requirements result in SAM artefacts that allow the full context of any data point to be captured using one common set of basic patterns.

This figure shows that the SAM editor tool is capable of:

- Creating a structure and naming/labeling the nodes
- Create labels in all languages

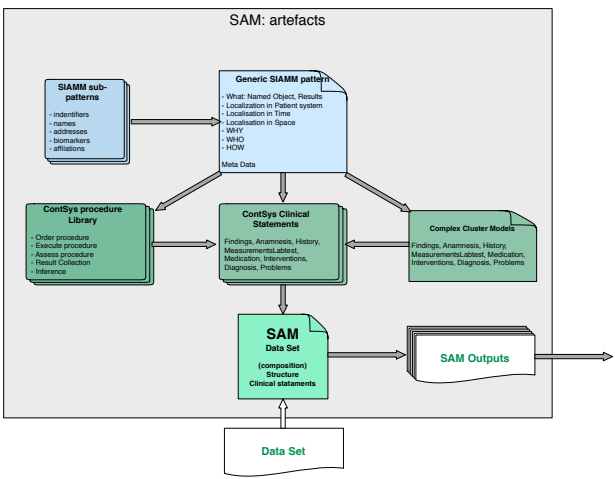
- Add necessary codes to labels of nodes
- Control the use of the structure via: cardinalities, occurrences and additional rules

The last feature is missing in Data dictionaries based on the ISO 11179 standard.

SAM Artefact Requirements:

1. *use one common pattern for any artefact reducing the degrees of freedom and allowing common shared basic patterns;*
2. *support concepts in the CEN/ISO System of Concepts for Continuity of Care, CEN General Purpose Information Components, CEN/ISO Health Information Service Architecture and ISO Subject of Care Identification.*

7.4. Tool: Document Manager content



The picture provides in detail an insight in the kinds of standardised building blocks that need to be managed in the Document Manager in order to be able to transform use cases and their data needs into a SAM. SAM's generate various outputs

The figure shows a screenshot of the deployed Document Manager.

The document manager allows:

- to input any document
- have complete version control over all documents
- provide meta-data about the content of the document, including codes for retrieval
- indicate what resources are associated with this document, such as: human and computer readable exchange formats, implementation guides and other supporting materials
- indicate what the relationship is between documents
- and add any comments by contributors

This tools allows the construction a collaborative environment for pools of contributors.

The depicted screenshots are from a Document Manager marketed by ERS and in use in the Ministry of Health in Spain.

7.5. Tool: SAM Data Dictionary

Data Dictionaries allow the definition of data elements that are part of datasets and their allowed attached codes.

A data dictionary[40], or metadata repository, as defined in the IBM Dictionary of Computing, is a "centralized repository of information about data such as meaning, relationships to other data, origin, usage, and format." The term may have one of several closely related meanings pertaining to databases and database management systems (DBMS).

The ISO/IEC 11179 - Metadata registries (MDR), addresses the semantics of data, the representation of data, and the registration of the descriptions of that data. It is through these descriptions that an accurate understanding of the semantics and a useful depiction of the data are found.

Data dictionaries are available in two variants: Passive and active Data Dictionaries.
Active Data Dictionaries allow on the fly translations of data elements and codes in one IT-system (data base) to data elements and codes in an other IT-system (database).

Several possible Data Dictionaries are available. Most are based on the ISO/IEC 11179 standard.

In the course of this project a data dictionary application has been supplied with a facility to import EN ISO 13606 archetypes that represent a SAM. As could be expected not all attributes in 13606 archetypes could be transformed to the 11179 application. In particular the ISO 11179 standard has weaknesses capturing structures. The ISO 11179 has facilities to express occurrences and cardinalities in text but not in a computer processable way.

The conclusion is that EN ISO 13606 archetypes/templates express better the datasets as modeled in SAM's.

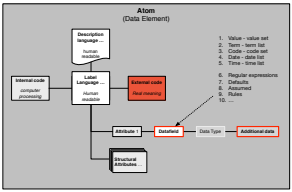
7.6. Tool: Terminology services

This subchapter corresponds to Part 3.1 of the project deliverables.
Based on desktop research this report will focus on tooling for terminological services. Other chapters deal with the deployment and governance. One Terminology server was used in the demonstrator.

SAM's represent the datasets as defined by healthcare providers and others. In order to make them more semantically interoperable codes are indispensable.

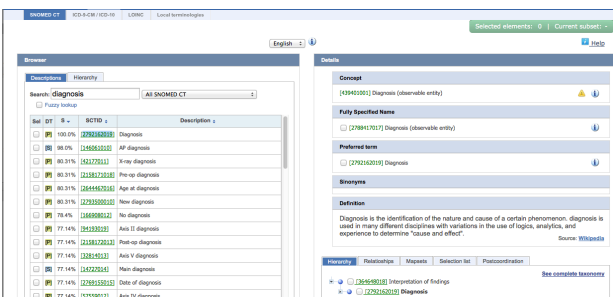
[40] http://en.wikipedia.org/wiki/Data_dictionary

Codes are needed to annotate the nodes in the SAM's and make them interoperable. In addition there is a need to add codes to results/data fields. Often these are lists of reserved words and lists with preselected codes that can be used as datum.



Terminology Servers are pieces of middle ware that are used to provide access to a coding systems from where one code can be selected, or from where a list with reserved terms can be selected to choose from, or a list of codes and descriptions that can be selected.

The Object Management Group (OMG) has published a standard interface for Common Terminology Services 2[41]. Such a middleware service can give access to a commercially available product where the vendor maintains the functionality and maintains the version of the codings systems that are needed. Not many products are available. During the present product we have made use of ITserver[42] by Indizen. One version is integrated with the SAM editor used in this project. There is a browser version, as well. This terminology service allows the manual selection from codes systems (SNOMED-CT, ICD9, ICD10, LOINC) and local terminologies, mappings between coding systems and the creation of subsets, and local value sets.



[41] http://www.omg.org/spec/CTS2/1.0/
http://www.ihtsdo.org/fileadmin/user_upload/doc/showcase/show13/SnomedCtShowcase2013_Abstract_23.pdf

[42] http://www.itserver.es/ITServer/common/index.faces

7.7. Tool: Technical Testing Service

This tool is mentioned in order to be complete.

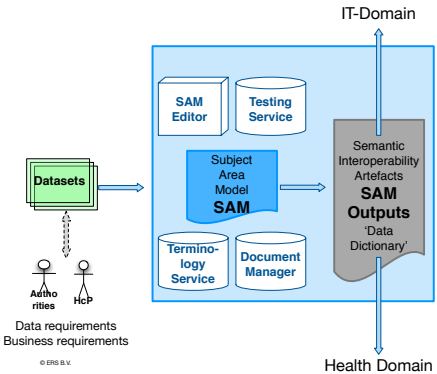
Much of the complex technical validation will rely on the parsers that are already in use in the other tools plus all computer processable outputs of the SAM's.

As part of the collaboration environment it would be nice if IT-software developers can test their implementations on-line.
Tools inside the HSE Tooling Environment make use of parsers that enforce that outputs are and stay conformant to the standards specifications. SAM and its outputs need human validation on clinical content only.

The Technical Testing Service and all parsers in the other tools will play an important role in the governance set-up.

7.8. Tool: Summary

- Tools are indispensable for the creation, maintenance, publication and support of the 'Normalised Data Platform'.
- Tool/services that are needed are: the SAM editor, a Document Manager, a Collaborative environments, Terminology Services and a testing Service to make available value sets and code sets
- Additional services that are needed are: licenses for the relevant coding systems such as: SNOMED-CT and LOINC plus all others that will be selected as next steps
- Additional in a next phase tools might be needed that help create Connectors for the implementation/integration at user sites


IT-Domain
SAM Editor
Testing Service
Datasets
Subject Area Model SAM
Semantic Interoperability Artefacts
SAM Outputs
'Data Dictionary'
Termino-logy Service
Document Manager
Autho rities
HCP
Data requirements
Business requirements
© ERS B.V.
Health Domain

8. Deployment / implementation

8.1. Deployment: Introduction

This chapter elaborates on next steps after the decision to embrace the proposed approach (IA-Rm and SAM's).

8.2. Deployment: Why is deployment important

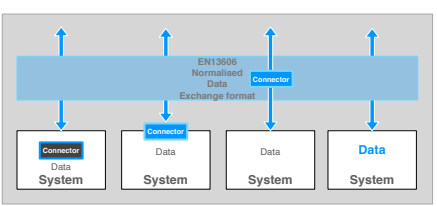
A big majority of Irish actors expressed that an evolutionary approach is called for, during workshops, that were held in Q3 and Q4 2013, and via the questionnaire.
Without a carefully planned set of next deployment steps the investment of the IA-RM, SAM's (EN ISO 13606 archetypes and templates as building blocks), the associated outputs and supporting services the benefits of the proposed national architecture can never be realised.
Expected benefits of the HSE Tooling Environment supported 'Normalised Data Platform' are:

- specifications of data needs readable by humans and computer processable by the IT-industry that can be used in procurement
- specifications technical outputs (archetypes) that facilitate a flexible integration of existing IT-solutions supporting national
- specifications technical outputs (archetypes) that when implemented support collaborative healthcare, flexible reporting, clinical auditing, clinical and pharmaceutical research
- specifications technical outputs (archetypes) that create a level playing field for existing IT-vendors and new companies (SME's) to create functionality based on normalised data

8.3. Deployment: Environment

Deployment in the market

The various actors in the IT-domain will use the IA-RM and the SAM's plus their associated outputs in the implementation process. The next figure depicts the different possible configurations:



The deployment steps must be designed not to disrupt the existing and new projects and solutions such as HealthLink and MidLis. Benefits of the IA-RM-SAM approach have been clearly identified by MedLis project management and a joint project with this clinical program is suggested as a stepwise approach to the involvement of other clinical program's.

The proposed deployment scenario is to allow existing exchange formats (HL7v2) to be used in existing situations. When data needs to be reported to receiving National repositories these repositories can absorb the data after a transformation process executed in an Integration Engine. Alternatively the senders can transform the data to the normalised format using Integration Engines.
Finally in the future there will be systems that internally will be able to process, exchange the data without the need for transformations.

The SAM's and associated outputs that will be published will constitute a 'Normalised Data Bus'.A set of domains in which outputs that help define and deploy the 'Normalised Data Bus 'are:

- **Requirements expression**: the formalised detailed requirements description of healths datasets, the SAM's
- **Procurement**: a series of SAM's and their associated outputs that can be used in procurement as a specification of the data needs of a system
- **Technical integration of existing solutions**: SAM's and their associated outputs that can be used in the integration of existing IT-systems using EN ISO 13606 connectors that deal with proprietary formats, HL7v2, HL7v3, HL7 CDA, Edifact, etc.
- **Technical integration for new solutions:** a series of SAM's and their associated outputs that can be used in procurement as a specification of the data needs to implemented in the interfaces and data base

A carefully designed set of next steps will allow an evolutionary deployment and wilf prevent investment losses.

Deployment at the National level

The deployment at the National level calls for one Governing organisation that is responsible and accountable for the Governance of the 'Normalised Data Bus' and all the needed resources.

8.4. Deployment: Summary

One of the requirements as expressed by the Workshops and Questionnaire is the need for an evolutionary approach.

Deployment of the 'Normalised Data Bus' is mostly in the domains of:

- **Requirements expression**: the formalised detailed requirements description of health datasets, the SAM's
- **Procurement**: a series of SAM's and their associated outputs that can be used in procurement as a specification of the data needs of a system
- **Technical integration of existing solutions**: SAM's and their associated outputs that can be used in the integration of existing IT-systems using EN ISO 13606 connectors that deal with proprietary formats, HL7v2, HL7v3, HL7 CDA, Edifact, etc.

- **Technical integration for new solutions:** a series of SAM's and their associated outputs that can be used in procurement as a specification of the data needs to implemented in the interfaces and data base

The four presented deployment possibilities support existing systems and their capabilities to be connected to the 'Normalised Data Bus' and slowly migrate overtime to the deployment of IT-systems that no longer need EN ISO 13606 connector.

9. Governance: Framework and Tooling

9.1. Governance: Introduction

Governance is defined as 'the way that a city, company, etc. is controlled by the people who run it.
This chapter will describe aspects of that what is needed to deploy the Information Architecture Reference Model (IA-RM), the Subject Area Models (SAM's) plus supporting services for production, testing, publishing and maintenance of all these artefacts.

9.2. Governance: Why is governance important

The IA-RM and SAM's plus supporting services create archetypes and templates exported in various formats and rely on supporting services. Without these outputs there will not be a resource shared by all. A semantic interoperability resource that creates the green level playing field for all actors, also known as the 'Normalised Virtual Data Bus'.
Without the 'Normalised Virtual Data Bus' full and safe semantic interoperability will be difficult or impossible. Without this shared resource exchange of data will be difficult, and thereby joint care, re-use of data for business management reporting, clinical reporting (referrals, discharge letters), ePrescribing, Personal Record, auditing, remuneration, and research.

9.3. Governance: Description

Governance is the set of measures taken by an organisation that functions in a controlled fashion.
This chapter will not go into details about what organisation will govern, its personal, political aspects, etc. At the organisational level ISO 9000[43] series of quality management standards will give guidance. There is a health specific equivalent[44].
This chapter will describe the specific governance aspects of the SAM's and their associated outputs.
The SAM's will be used by healthcare actors to define their data needs for systems and exchange between systems; and they will be used by the IT-industry to comply with.
The complete life cycle of all the artefacts that are produced need to be covered.
Artefacts are produced, tested/validated, placed in the market/published, maintained and finally deprecated/removed. IT-vendors and or data feeder systems will claim compliance with the pub-

[43] http://www.iso.org/iso/home/standards/management-standards/iso_9000.htm

[44] ISO 13485:2012 is the medical industry's equivalent of ISO 9001:2008. Whereas the standards it replaces were interpretations of how to apply ISO 9001 and ISO 9002 to medical devices, ISO 13485:2003 is a stand-alone standard. Because ISO 13485 is relevant to medical devices manufacturers (unlike ISO 9001, which is applicable to any industry), and because of the differences between the two standards relating to continual improvement, compliance with ISO 13485 does not necessarily mean compliance with ISO 9001:2008 (and vice versa)

lished requirements and need qualification or certification that their products are in conformance to the SAM's and their associated outputs.

The Medical Device Directive (MDD)[45] is an example of how, on the European level quality aspects of artefacts, used in healthcare, are managed. Recently the scope of the MDD is extended to software. Essential in this approach is the identification of risks and the mitigation of it by means of technical solutions, workarounds or warnings.
ISO 13485:2012 is the medical industry's equivalent of ISO 9001:2008. Whereas the standards it replaces were interpretations of how to apply ISO 9001 and ISO 9002 to medical devices, ISO 13485:2003 is a stand-alone standard. Because ISO 13485 is relevant to medical devices manufacturers (unlike ISO 9001, which is applicable to any industry), and because of the differences between the two standards relating to continual improvement, compliance with ISO 13485 does not necessarily mean compliance with ISO 9001:2008 (and vice versa).

9.3.a. Validation, Testing and certifying of SAM's

The ANTILOPE project[46] drives eHealth interoperability in Europe and beyond. Between 2013 and 2015 key national and international organisations will work together to select and define eHealth standards and specifications. They will create, validate and disseminate a common approach for testing and certification of eHealth solutions and services in Europe. They focus on testing of 'profiles' and certification of conformance claims to standards in the interoperability domain. 'Profiles' in the context of ANTILOPE and IHE[47] (Integrating the Healthcare Enterprise) is an organisation of vendors that, starting on a use case create a 'profile' of existing messaging standards as detailed implementation document. Mostly these messaging standards are HL7v2 or CDA.

At the European level there is the European Interoperability Framework[48] (EIF). A set of recommendations which specify how Administrations, Businesses and Citizens communicate with each other within the EU and across Member States borders. eHealth is part of it. Projects are underway that at the European level deal with the sharing of re-usable Semantic Interoperability artefacts and quality assurance of these artefacts.

In the case of the IA-RM and SAM's the SAM's are defined using the CEN/ISO EHR-communication standard that do NOT need profiles the way IHE is producing them, because these SAM's are produced using the Two-Level-Modeling Paradigm.

[45] http://ec.europa.eu/health/medical-devices/documents/revision/index_en.htm

[46] http://www.iso.org/iso/home/standards/management-standards/iso_9000.htm

[47] http://www.ihe.net

[48] http://en.wikipedia.org/wiki/European_Interoperability_Framework

Health and clinical validation

Important aspects of any artefact are on one hand the functional requirements (the datasets and the codes used) and the technical requirements. On the other hand there are technical requirements. In the case of SAM's based on the EN ISO 13606 standard these artefacts representing the (health) data needs need to be validated by three or more (clinical domain) specialists by inspection and comparison of the provided and agreed dataset and the SAM's.

Other artefacts that are defined and used are value- and code lists that are made available via a Terminology Service conforming to the OMG CTS2 standard[49].

Technical validation

The technical expressions that the IT-industry use, do not need to be validated because of the fact that they are produced using tools based on the Two-Level-Modeling-Paradigm. By definition all artefacts are conformant to the reference models and data specifications and can when needed be tested by the various parsers that were used to generate them.

Product validation

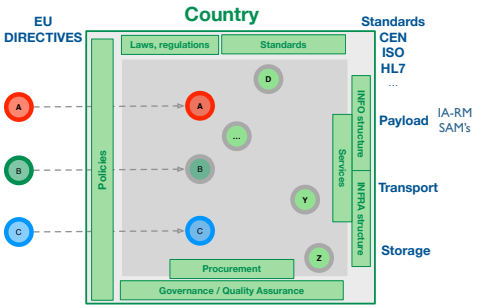
IT-vendors will claim conformance to the SAM's. The ANTILOPE project documents describe the way in which vendors can be tested and qualified cq certified.
Products that are conformant to the requirements can only be tested under laboratory testing conditions.

Implementation validation

Products that are conformant under laboratory testing conditions need to be tested/validated in real life conditions after implementation of the software and all SAM's and their associated outputs. The ANTILOPE project has produced documents that can be used under these conditions.

[49] http://www.omg.org/spec/CTS2/1.1/

9.4. Governance: Summary



The 'Normalised Virtual Data Bus' (an INFOStructure as defined and made possible by: IA-RM, SAM's and associated outputs plus services) needs to be governed.

In order to have a flexible and patient safe exchange of health data between IT-systems several aspects need to be addressed, and governed:

- Governmental policies, rules and regulations
- Policies based on open international standards
- Standards that specify an INFOstructure for the payload based on the IA-RM/SAM's and supporting services
- Standards that specify an INFRAstructure that allows the safe transport, storage and security aspects plus supporting services
- Quality assurance/validation/testing of hard and software

Procurement of the qualified/certified hard- and software

Several stakeholders will be involved in their natural role:

Political/organisational

It must be a political decision to appoint/erect an organisation that is responsible and accountable for the organisation that governs the eHealth INFOstructure in all its aspects.
A Quality Management System must be in place.

One or more laws plus supporting rules and regulations will be needed dealing with subjects such as: unique identification of persons (patients, healthcare providers) and relevant organisations, the personal privacy, rules for procurement of health-IT systems, quality assurance of healthcare processes, etc.
it is at the level of the Government that an active participation in International, European, organisations will be needed at various levels.

Standard Developing Organisations
The IA-RM and SAM's will be produced on the basis of open International standards such as CEN, ISO, IHTSDO and the WHO. At the technical infrastructure level standards will play an important role, also.
Representatives from Ireland will have to actively take part in these organisations.

Info- and Infrastructure services
At least one organisation will have to be responsible for the creation and maintenance of the SAM's and supporting services such as: Document Management, Terminology services, authorisation of healthcare providers, technical exchange of messages, etc.

Health domain
Groups of healthcare providers and (governmental) bodies have data needs expressed as datasets. These groups provide input and help validate the SAM's and associated outputs that are produced and maintained by an health informatician.
Possibly the organisation for the Info- and Infrastructure will interacts with the professional bodies

Health informatics
Health informaticians need to support the groups and organisations that express data needs and validate the SAM's, and outputs. These experts must be able to use the software tools, have knowledge about the health domain, about the standards involved, including how to produce archetypes, and how to use codes from the various coding systems in use.

Possibly the Health Informaticians will be specifically IT and Semantic Interoperability trained health professionals such as MD's, and Nurses.

Technical (IT)
The organisation will maintain an IT related series of services to support the primary process that is executed by the organisation for Info- and Infrastructure services.
E.g. Website, Collaboration environment, Document handler, terminology/ value -set server) Catalogue of deployed Standards.

10. Catalogue of deployed Standards

Standard name	Description	Comments
CEN 14822	General Information Components	3 part standard specifying the European information needs in healthcare based on 15 years of message production
CEN/ISO 12967	Health Information Service Architecture	Specification that defines the interfaces in an EHR-system
EN ISO 13606:2008	EHR Communication	5 part standard specifying the exchange of any partial or complete EHR
CEN/ISO-13940	System of concepts for continuity of care	Terminological standard about processes in health and care
CEN/TS 14796	CEN Data Types	A CEN data type standard used in the present EN ISO 13606. There is consensus in the standards community to represent these data types as a profile of the ISO 21090. To be replaced by a profile of ISO 21090:2011 Harmonized data types for information interchange.
GS1	Industry standards to identify products and characteristics.	Industry standards to identify products and characteristics.
HL7 v2.x	Health Level Seven V2.x Messaging standard	The HL7 version 2 standard (also known as Pipehat) has the aim to support hospital workflows. It was originally created in 1989 Version 2 defines a series of electronic messages to support administrative, logistical, financial as well as clinical processes. Since 1987 the standard has been updated regularly, resulting in versions 2.1, 2.2, 2.3, 2.3.1, 2.4, 2.5, 2.5.1 and 2.6. The v2.x standards are backward compatible (e.g., a message based on version 2.3 will be understood by an application that supports version 2.6).
HL7 v3 ISO/HL7 21731	Health Level Seven v3 Message standard	The HL7 version 3 standard has the aim to support all healthcare workflows. Development of version 3 started around 1995, resulting in an initial standard publication in 2005. The v3 standard, as opposed to version 2, is based on a formal methodology (the HDF) and object-oriented principles. The HL7 version 3 standard has the aim to support all healthcare workflows. Development of version 3 started around 1995, resulting in an initial standard publication in 2005. The v3 standard, as opposed to version 2, is based on a formal methodology (the HDF) and object-oriented principles. The Reference Information Model[19] (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.

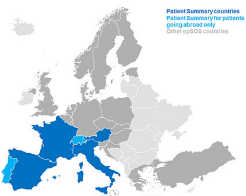
Standard name	Description	Comments
HL7v3 CDA	Health Level Seven Clinical Document Standard	The HL7 Clinical Document Architecture (CDA) is a XML-based markup standard intended to specify the encoding, structure and semantics of clinical documents for exchange. CDA is an ANSI-certified standard from Health Level Seven (HL7.org). Release 1.0 was published in November, 2000 and Release 2.0 was published with the HL7 2005 Normative Edition. CDA specifies the syntax and supplies a framework for specifying the full semantics of a clinical document. It defines a clinical document as having the following six characteristics: Persistence, Stewardship, Potential for authentication, Context, Wholeness, Human readability.
IHTSDO - SNOMED-CT	A huge collection of clinical concepts	A huge collection of clinical concepts IHTSDO co-operates with the WHO on ICD-11. Not all parts of SNOMED-CT are safe to use.
ISO 13119	Clinical Knowledge Resources - Metadata	
ISO 13485	Medical devices - Quality management systems	
ISO 18308	Requirements for an Electronic Health Record Architecture	
ISO 21090: 2011	Health informatics -- Harmonized data types for information interchange	This specification is too complex to be used as is. A profile for specific use is necessary. For use as part of EN ISO 13606 a profile will be published in the present 13606 renewal phase.
ISO 22220	Identification of Subjects of Care	Demographics (identification, names, addresses, other addresses, gender, etc.)
ISO 3166	International Standard for country codes	
ISO 639	Nomenclature used to classify all known languages.	Nomenclature used to classify all known languages.
ISO 639	Codes for the representation of names of languages	
ISO 8601	Representation of dates and times	
ISO 9000	ISO 9000 - Quality management series	The ISO 9000 family addresses various aspects of quality management and contains some of ISO's best known standards. The standards provide guidance and tools for companies and organizations who want to ensure that their products and services consistently meet customer's requirements, and that quality is consistently improved.
ISO/IEC 11179-3	Metadata registries (MD): Registry metamodel and basic attributes	Specifies the functionality of a data dictionary

Standard name	Description	Comments
ISO/IEC 17011	Conformity assessment – General requirements for accreditation	
ISO/IEC 17020	Conformity assessment – Requirements for the operation	
ISO/IEC10746-1	RM/ODP	Open Distributed Processing - Reference Model: Overview
LOINC	Logical Observation Identifiers Names and Codes huge collection of clinical concepts	Database and universal standard for identifying medical laboratory observations
MedDRA	Medical Dictionary for Regulatory Activity	In developing and continuously maintaining MedDRA, ICH endeavors to provide a single standardised international medical terminology which can be used for regulatory communication and evaluation of data pertaining to medicinal products for human use. As a result, MedDRA is designed for use in the registration, documentation and safety monitoring of medicinal products through all phases of the development cycle (i.e., from clinical trials to post-marketing surveillance).
OMG CTS2 version1.1	Common Terminology Services 2	http://www.omg.org/spec/CTS2/1.0/
UCUM	Unified Code for Units of Measure	The Unified Code for Units of Measure is a code system intended to include all units of measures being contemporarily used in international science, engineering, and business. http://unitsofmeasure.org/ucum.html http://unitsofmeasure.org/trac/
WHI ICHI	International Classification of Health Interventions (ICHI)	The purpose of this classification is to provide Member States, health care service providers and organizers, and researchers with a common tool for reporting and analysing the distribution and evolution of health interventions for statistical purposes. It is structured with various degrees of specificity for use at the different levels of the health systems, and uses a common accepted terminology in order to permit comparison of data between countries and services
WHO ATC/DDD SYSTEM	The Anatomical Therapeutic Chemical (ATC) classification system and the Defined Daily Dose (DDD)	The Anatomical Therapeutic Chemical (ATC) classification system and the Defined Daily Dose (DDD) as a measuring unit are recommended by the WHO for drug utilization studies. The system is widely used internationally and the number of users is increasing

Standard name	Description	Comments
WHO ICD, x	The International Classification of Diseases (ICD)	The International Classification of Diseases (ICD) is the standard diagnostic tool for epidemiology, health management and clinical purposes. This includes the analysis of the general health situation of population groups. It is used to monitor the incidence and prevalence of diseases and other health problems. It is used to classify diseases and other health problems recorded on many types of health and vital records including death certificates and health records. In addition to enabling the storage and retrieval of diagnostic information for clinical, epidemiological and quality purposes, these records also provide the basis for the compilation of national mortality and morbidity statistics by WHO Member States. It is used for reimbursement and resource allocation decision-making by countries. ICD-10 was endorsed by the Forty-third World Health Assembly in May 1990 and came into use in WHO Member States as from 1994. The 11th revision of the classification has already started and will continue until 2017.
WHO ICF	The International Classification of Functioning, Disability and Health, known more commonly (ICF)	The International Classification of Functioning, Disability and Health, known more commonly as ICF, is a classification of health and health-related domains. As the functioning and disability of an individual occurs in a context, ICF also includes a list of environmental factors. ICF is the WHO framework for measuring health and disability at both individual and population levels. ICF was officially endorsed by all 191 WHO Member States in the Fifty-fourth World Health Assembly on 22 May 2001(resolution WHA 54.21) as the international standard to describe and measure health and disability.
WHO ICPC	International Classification of Primary Care, Second edition (ICPC-2)	WHO has accepted ICPC-2 within the WHO FIC mainly as a reason for encounter classification, and users may use it as a classification for primary care or general practice wherever applicable. ICPC-2 classifies patient data and clinical activity in the domains of General/Family Practice and primary care, taking into account the frequency distribution of problems seen in these domains. It allows classification of the patient's reason for encounter (RFE), the problems/diagnosis managed, interventions, and the ordering of these data in an episode of care structure
WHO INN	International Nonproprietary Names (INN)	International Nonproprietary Names (INN) facilitate the identification of pharmaceutical substances or active pharmaceutical ingredients. Each INN is a unique name that is globally recognized and is public property. A nonproprietary name is also known as a generic name

11. Introduction: epSOS

epSOS aims to design, build and evaluate a service infrastructure that demonstrates cross-border interoperability between electronic health record systems in Europe.



Cross-border eHealth Services

epSOS attempts to offer seamless healthcare to European citizens. Key goals are to improve the quality and safety of healthcare for citizens when travelling to another European country. Moreover, it concentrates on developing a practical eHealth framework and ICT infrastructure that enables secure access to patient health information among different European healthcare systems. epSOS can make a significant contribution to patient safety by reducing the frequency of medical errors and by providing quick access to documentation as well as by increasing accessibility of ones prescribed medicine also abroad. In emergency situations, this documentation provides the medical personnel with life-saving information and reduces the (sometimes needless) repetition of diagnostic procedures.

The technical, legal and organizational concepts developed within the framework of the project are subject to an extensive practical testing phase which will last until the end of the project. epSOS will tests cross-border eHealth services in the following areas.

In a first phase:

- *Patient Summary: access to important medical data for patient treatment.*
- *Cross-border use of electronic prescriptions ("ePrescription" - or "eMedication" systems)*

In the extended project phase - which started 2011- the epSOS project team consolidates, scales up and operationalizes the epSOS Services for ID management, security, semantics and standards. Additional epSOS Services like the access of patients to their data or the Medication Related Overview (MRO) are analyzed and will be tested if feasible.

For the first time, patients in Europe have the opportunity to use cross-border eHealth services when seeking healthcare in participating epSOS pilot countries - whether as tourists, business travelers, commuters or exchange students.

11.1. epSOS developments

EpSOS is developed during two EU-projects and is the basis for an European Guideline on the Patient Summary[50].

Two Use Cases were used to guide developments:

- *USE CASE 1: The patient is an occasional visitor to the country of treatment, for example someone on holiday or attending a business meeting. The distinguishing characteristic is that this type of visit is irregular, infrequent and may not be repeated. This is a type of incidental encounter where the health professional will not normally have a previous record of the person seeking care and where the health professional does not know the patient.*
- *USE CASE 2: The patient is a regular visitor to another country from his or her country of origin, for example someone who lives in one country but works in another. The distinguishing characteristic is that this type of visit is regular, frequent and the person seeking care may be accustomed to using services in the country where he or she works as a matter of personal convenience. In this situation, the health professional may have some information available from previous encounters; the patient may therefore have a patient record locally stored in country B and possibly also a PS in country A, and both sources of information could be consulted.*

Three concepts were defined:

1. *The **Patient Summary** is an identifiable "dataset of essential and understandable health information" that is made available "at the point of care to deliver safe patient care during unscheduled care [and planned care] with its maximal impact in unscheduled care"; it can also be defined at a high level as: "the minimum set of information needed to assure healthcare coordination and the continuity of care".*
2. *The **basic dataset** is defined as a set of essential health information that needs to be sent from a clinical point of view in order to be able to deliver safe care to the patient (focused in unscheduled care). The information of the basic dataset must always be available.*
3. *The **extended dataset** is defined as the minimum amount of recommended health information from a clinical point of view that needs to be exchanged between Member States. These fields should be completed whenever possible.*

EU-Member States are fully free to implement the dataset or not using any technical standard according to this EU Recommendation/Guideline. Some will implement it using the HL7 CDA standard, other are allowed to opt for a solution based on the EN ISO 13606 EHR-communication standard.

[50] Guidelines on minimum/non-exhautice patient summary dataset for electronic exchange in accordance with the cross-border Directive 2011/24/EU Release 1.As adopted by the eHealth Network (19-11-2013) http://ec.europa.eu/health/ehealth/docs/guidelines_patient_summary_en.pdf

Datasets use coding systems. Using a set of agreed requirements coding systems were selected for each data element in the Patient Summary. These coding systems are collected in the Master Value Sets Catalogue (MVC). Because Europe has 23 languages and that the Patient Summary must be used for cross-border exchanges a Master Translation/Transcoding Catalogue (MTC) is published.

[xxx References to these services]

11.2. epSOS: Infrastructure

The epSOS project developed a set of profiles in order to be able to request en send the needed Patient Summary. This infrastructural aspect is outside of the scope of this project but will be described briefly.

11.3. SAM: epSOS dataset

The Subject Area Model chosen for this project is the epSOS dataset. In this chapter this dataset is presented.

It has three main components:

- epSOS Patient Data with data needed to identify the patient.
- epSOS Patient Summary data that defines data about the summary document
- epSOS Patient Clinical Data. This chapter reflects the consensus between European Member States in the epSOS project about the clinical content that is exchanged with the epSOS patient summary.

The epSOS data-set consist of two related aspects:

- Structure consisting of connected named and coded nodes;
- Per node coded named attributes and code sets;
- End nodes carry the results as data points. This recorded data can be numeric, text, but codes and code lists and value sets.

epSOS has defined its collection of codes, codes sets and value sets needed to populate the structure and give it common, shared, unique meaning.

11.4. SAM: Coding systems as used by epSOS

The Subject Area Model based on epSOS will have to describe the relevant codes and coding systems and value sets before it can be used for semantic interoperability.
epSOS describes the decision taken in the epSOS Master Value Set Catalogue (MCV[51]).

[51] epSOS D3.5.2 Appendix D epSOS Master Value Set Catalogue v0.0.3 2010215.doc

The next table defines the epSOS52 dataset in detail.

The next chapter will describe the semantic interoperability artefacts that will be used in the Proof of Concept of this HSE project and that are based on this EU-Guideline for the Patient Summary.

[52] Guidelines on Minimum/Non-exhaustive Patient Summary Dataset for Electronic Exchange in Accordance with the Cross-Border Directive 2011/24EU. Release 1. DD 191-11-2013

PATIENT ADMINISTRATIVE DATA				
Variable (nesting level 1)	**Variables (nesting level 2)**	**Variables (nesting level 3)**	**DEFINITION AND COMMENTS**	**BASIC (Basic)/ EXTENDED (Ext) DATASET**
Identification[1]	National healthcare patient ID	National healthcare patient ID	Country ID, unique to the patient in that country. Example: ID for United Kingdom patient	Basic
Personal information	Full name	Given name	The first name of the patient (example: John). This field can contain more than one element.	Basic
		Family name/surname	This field can contain more than one element. Example: Español Smith Note: some countries require surnames to be the birth name [to avoid potential problems with married women's surnames).	Basic
	Date of birth	Date of birth	This field may contain only the year if the day and month are not available, e.g. 01/01/2009	Basic
	Gender	Gender code	This field must contain a recognized valid value.	Basic
Contact information	Address[2]	Street	Example: Oxford Street	Ext
		House number	Example: 221	Ext
		City	Example: London	Ext
		Post code	Example: W1W 8LG	Ext
		State or province	Example: London	Ext
		Country	Example: UK	Ext
	Telephone no.	Telephone no.	Example: +45 20 7025 6161	Ext
	e-mail	e-mail	Example: jens@hotmail.com	Ext
	Preferred HP/HPO to contact[3]	Name of the HP/HPO	Name of the HP/ HPO that has been treating the patient. If this is an HP, the structure of the name will be the same as described in 'Full name' (given name, family name/surname).	Basic
		Telephone no.	Example: +45 20 7025 6161	Basic
		e-mail	e-mail of the HP/legal organization	Basic
	Contact person/ legal guardian (if available)	Role of that person	Legal guardian or contact person	Ext
		Given name	The first name of the contact person/guardian (example: Peter). This field can contain more than one element.	Ext
		Family name/surname	This field can contain more than one element. Example: Español Smith	Ext
		Telephone no.	Example: +45 20 7025 6161	Ext
		e-mail	e-mail of the contact person/legal guardian	Ext
Insurance information	Insurance number	Insurance number	Example: QQ 12 34 56 A	Ext

[1] Dataset that enables the univocal identification of the patient
[2] May vary by country
[3] A health professional in country A may need a contact (health professional/healthcare provider) who knows the patient.

PATIENT CLINICAL DATA				
Variable (nesting level 1)	**Variables (nesting level 2)**	**Variables (nesting level 3)**	**DEFINITION AND COMMENTS**	**BASIC (Basic)/ EXTENDED (Ext) DATASET**
Alerts	Allergy	Allergy description	Description of the clinical manifestation of the allergic reaction. Example: anaphylactic shock, angioedema (the clinical manifestation also gives information about the severity of the observed reaction)	Basic
		Allergy description ID code	Normalized identifier	Basic
		Onset date	Date of the observation of the reaction	Ext
		Agent	Describes the agent (drug, food, chemical agent, etc.) that is responsible for the adverse reaction	Basic
		Agent ID code	Normalized identifier	Basic
	Medical alert information (other alerts not included in allergies)	Healthcare alert description	Medical alert information: any other clinical information that is essential to know so that the life or health of the patient does not come under threat. Example 1: Intolerance to aspirin due to gastrointestinal bleeding. Example 2: intolerance to captopril because of cough (the patient is not allergic but cannot tolerate it because of persistent cough).	Basic
		Healthcare alert ID code	Normalized identifier	Basic
Medical history	Vaccinations	Vaccinations	Contains each disease against which the patient has been immunized	Ext
		Brand name		Ext
		Vaccination ID code	Normalized identifier	Ext
		Vaccination date	Date when the immunization was given	Ext
	List of resolved, closed or inactive problems	Problem description	Problems or diagnoses not included in the definition of "current problems or diagnosis". Example: hepatic cyst (the patient has been treated with an hepatic cystectomy that solved the problem, which is therefore a closed problem)	Ext
		Problem ID code	Normalized identifier	Ext
		Onset time	Date of onset of problem	Ext
		End date	Problem resolution date	Ext
		Resolution circumstances	Describes the reason for which the status of the problem changed from current to inactive (e.g. surgical procedure, medical treatment, etc.). This field includes "free text" if the resolution circumstances are not already included in other fields such as surgical procedure, medical device, etc., e.g. hepatic cystectomy (this will be the resolution circumstances for the problem "hepatic cyst" and will be included in surgical procedures).	Ext
	Surgical procedures prior to the past six months	Procedure description	Describes the type of procedure	Basic
		Procedure ID (code)	Normalized identifier	Basic
		Procedure date	Date when procedure was performed	Basic

Variable (nesting level 1)	Variables (nesting level 2)	Variables (nesting level 3)	DEFINITION AND COMMENTS	BASIC (Basic)/ EXTENDED (Ext) DATASET
Social history	Social history observations	Social history observations related to smoking, alcohol and diet	Health-related "lifestyle factors" or "lifestyle observations" Example: cigarette smoker, alcohol consumption	Ext
		Reference date range	Example: from 1974 to 2004	Ext
Pregnancy history	Expected date of delivery	Expected date of delivery	Date on which the woman is due to give birth. Year, month and day are required (e.g. 01/01/2014).	Ext
Physical findings	Vital signs observations	Blood pressure	One blood pressure value, which includes systolic blood pressure and diastolic blood pressure	Ext
		Date when blood pressure was measured	Date when blood pressure was measured	Ext
Diagnostic tests	Blood group	Result of blood group	Result of blood group test performed on the patient	Ext
		Date	Date on which the blood group test was performed. This field may contain only the year if the day and month are not available (e.g. 01/01/2009).	Ext

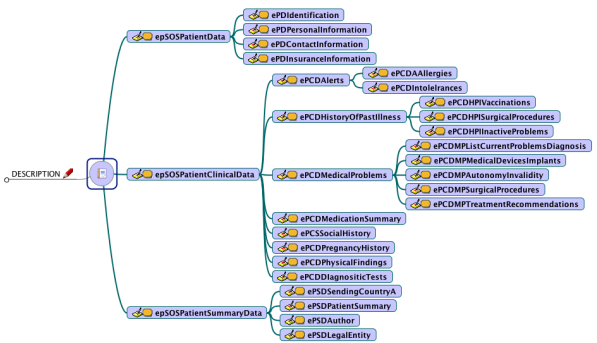
PATIENT ADMINISTRATIVE DATA				
Variable (nesting level 1)	Variables (nesting level 2)	Variables (nesting level 3)	DEFINITION AND COMMENTS	BASIC (Basic)/ EXTENDED (Ext) DATASET
Country	Country	Country	Name of country A	Basic
Patient Summary	Date created	Date created	Date on which PS was generated	Basic
	Date of last update	Date of last update	Date on which PS was updated (date of most recent version)	Basic
Nature of the PS	Nature of the PS	Nature of the PS	Defines the context in which it was generated. Distinguishes between three methodological approaches for generating the PS: direct human intervention by an HP, automatically generated approach and mixed approach	Basic
Author organization	Author organization	Author organization	At least one author organization (HCP) shall be listed. If there is no HCP, at least one HP shall be listed.	Basic

11.5. SAM: Semantic Interoperability Artefacts Modeling Method

The epSOS dataset will be represented in this project for the proof of concept using a method developed by ERS and published by the EN ISO 13606 Association for the production of semantic interoperability artefacts called archetypes. The name of this method is the Semantic Interoperability Artefact Modeling Method (SIAMM). This method is used in EU projects (SemanticHealthNet, SALUS, and soon in EXPAND), in addition it is a candidate for inclusion in the EN ISO 13606 EHR communication standard during its renewal process.

The goal of this method is to design archetype patterns that allow to record all of the semantics, the information around data points such that these data points describe fully all that is necessary to interpret the data safely.

11.5.a. SIAMM General Pattern



One generic SIAMM archetype pattern is used to derive in a structured way all other re-usable patterns. In addition SIAMM is (and will be) aligned with two other important standards: CEN/ISO System of Concepts for Continuity of Care and CEN/ISO Health Information Service Architecture.

The generic SIAMM pattern describes for each clinical data point (clinical statement) the complete context:

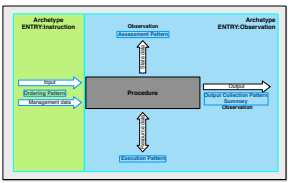
WHAT: The name of the subject of the clinical statement and the associated result(s). E.g. 'Diagnosis is Diabetes with code XYZ', 'Blood Glucose measurement result = 7.4 mM/L', 'Order the execution of procedure ZYX'.

- WHERE: features to locate the topic in time, place and when relevant the patient system.
- WHO: The items, persons, or organisations that participate in the clinical statement.
- WHY: The reasons why this clinical statement is documented in the EHR.

- HOW: The circumstances that are necessary to interpret the clinical statement correctly and safely. E.g. 'Body Weight measured without clothing'.' Systolic Blood Pressure measured while in rest'.

11.5.b. SIAMM Specialised Sub-Patterns

In contrast to other methods of archetype modeling, SIAMM is aligned with the CEN/ISO Standard Concepts for Continuity of Care. One of the consequences is that what is modeled is modeled from the perspective of processes (Pathways, Protocols, Tasks).



One single general SIAMM pattern is specialised into sub-patterns that define what is documented about:

Ordering,

- Execution,
- Assessment and
- Summary after completion plus
- Inferences.

In the case of the epSOS patient summary production in an EHR-system a Procedure to create the epSOS Patient Summary is ordered, executed and when finished there is an epSOS Patient summary.

The data in the Patient Summary in the sending system is re-use of existing data and not de novo data generated by the author during a patient contact. The SIAMM artefacts will reflect this because one of the attributes in the MetaData of the Artefact indicates that it is not de novo data -as the result of a diagnostic or therapeutic process- but an administrative process.
The epSOS Patient Summary that will be created by the Procedure is a Composition with attributes that this is according to ContSys not 'non-ratified clinical data', since it is duplicated data for the purpose of reporting. Observe that the document/message itself can and must be signed (ratified) as document.
The data in the Composition will not be searched since it is 'not-ratified', not active.

In the receiving system a choice by the receiving Healthcare Provider must be made to designate the data from the epSOS Patient Summary as data that will or will not become searchable in his EHR-system by inserting it in the health record as non-ratified clinical data or 'normal' clinical data.

The table lists the mapping of the epSOS dataset on the epSOS Template and what SIAMM archetypes will be used to build this Template. The SIAMM archetypes that will be used to populate the epSOS Template will be ENTRY archetypes, that use one of the sub-patterns that model one of the stages in the life-cycle of any Procedure.
E.g.

- ENTRY:ReportingItemProcedure:Allergies, or
- *ENTRY:ReportingItemProcedure:MedicinalProduct, etc*.

All using the SIAMM Summary sub-pattern in order to indicate that it is the result after the execution of a Procedure that will query the existing data and insert it into the epSOS Patient Summary Template (Composition).

These SIAMM archetypes will be enriched by SNOMED-CT codes and used to populate the HSE SAM Library and HSE Data Dictionary.

Layer 1	Layer 2	Layer 3	Mapping to 13606 and ContSys	Specialisation hierargy	Specialisation hierargy	Specialisation hierargy
epSOS exchange artefact - Template			13606: EHR-extract ContSys:electronic patient summary	electronic health record extract		
epSOS Dataset			13606: Composition ContSys:electronic record compo-nent (s)			
epSOSPatientData			13606: Section ContSys: subject of care specialisation of: healthcare actor and person role			
	ePDIdentification		13606: ENTRY: via NamedObject	healthcare actor		
	ePDPersonalInformation					
	ePDContactInformation					
	ePDInsuranceInforma-tion					
epSOSPatientClini-calData			13606: Section ContSys: observed condition (s)			
	ePCDAlerts		13606: Section ContSys: risk condition (s)	potential health condition	health condi-tion	health issue
		ePCDAAllergies	13606: ENTRY ContSys: risk condition (s)	potential health condition	health condi-tion	health issue
		ePCDAIntoller-ances	13606: ENTRY ContSys: risk condition (s)	potential health condition	health condi-tion	health issue
	ePCDIHistoryOfPastIll-ness		13606: Section ContSys:health condition (s) Specialisation of: health issue			
		ePCDHPIVaccina-tions	13606: ENTRY ContSys: healthcare treatment	(healthcare activity element)	healthcare activity	
		ePCDHPISurgical-Procedures	13606: ENTRY ContSys: healthcare treatment	(healthcare activity element)	healthcare activity	
		epCDHPIInac-tiveProblems	13606: ENTRY ContSys: health problem	potential health state	health condi-tion	health issue
	ePCDMedicalProblems		13606: Section ContSys: health problem list Specialisation of health condition			
		ePCDMPListCur-rentProblemsDiag-nosis	13606: ENTRY ContSys: health problem	health condition	health state	
		epPCDMPMedi-calDevicesImplants	13606: ENTRY ContSys: healthcare treatment (medical device)	(healthcare activity element)	healthcare activity	
		epPCDMPAutono-myInvalidity	13606: ENTRY ContSys: observed condition		health condi-tion	health issue
		ePCDMPSurgical-Procedures	13606: ENTRY ContSys:healthcare treatment (surgery procedure)	(healthcare activity element)	healthcare activity	
		epCDMPTreatmen-tRecommendations	13606: Section (or Folder) Contsys: care plan Generalisation of: uniprofesional care plan, or multi-professional care plan Aggregation of: care plan, healthcare activity, healthcare activities bundle			care plan
			13606:ENTRY Any: healthcare activity, other sub-plan	(healthcare activity element)	healthcare activity	

Layer 1	Layer 2	Layer 3	Mapping to 13606 and ContSys	Specialisation hierargy	Specialisation hierargy	Specialisation hierargy
	epSOSPatientClinicalData (continued)			13606: Section ContSys: observed condition(s)		
		epCDMedicalSummary	13606: ENTRY ContSys: observed condition	health condition	health issue	
		epCDSocialHistory	13606: ENTRY ContSys: observed condition	health condition	health issue	
		epCDPregnancyHistory	13606: ENTRY ContSys: observed condition	health condition	health issue	
		epCDPhysicalFindings	13606: ENTRY ContSys: observed condition	health condition	health issue	
		epCDDiagnosticTests	13606: ENTRY ContSys: healthcare investigation	health activity element	healthcare activity	
	epSOSPatientSummaryData			13505: Section		
		epPSDPatientSummary	DateCreated Date Last Update	13606: ENTRY: ??? ContSys: ???		
		epPSDAuthor		13606: ENTRY associated with ContSys: healthcare provider, healthcare organisation		
		epPSDLegalEntity				
		epPSDSCountry		sending country A 13606: ENTRY		

11.6.SAM: Semantic Interoperability: Artefacts and Coding Systems

This chapter describes coding systems that can be used as part of the Subject Area Models (SAM's).

11.6.a. Introduction: Codes, Code sets and value sets

Human text is many time arbitrary and fuzzy. In order to reduce fuzziness in communication eHealth systems make use of agreed code sets. These codes are like words in a dictionary and describe in detail the meaning of the word, its preferred spelling, etc.

A code is a combination of a unique number, a label used for presentation and a description. Together with the Coding System it stems from the code and its meaning can be described fully.

E.g.
Coding System = QCODE, version 2008
Unique Number (code) = 12xyz12
Label = Aorta
Description: The aorta (/eɪˈɔrtə/) is the largest artery in the human body, originating from the left ventricle of the heart and extending down to the abdomen, where it bifurcates into two smaller arteries (the common iliac arteries).

Sometimes only selected (Non-coded) text or numbers are allowed in a data point. These are expressed as Value-sets.

E.g.
Allowed text that can be entered in a data point is: Male, Female, Unresolved.
Allowed numbers in a data point are: 1, 2, 3, 4, and 5. or 1940, 1950 or 1960.

11.6.b. Why are coding systems important

User groups documenting the provision of healthcare in one location and/or in one clinical domain share the meaning of words easily because of many years of education and training.
The IT-systems deal can with words/text.
When communication with the 'outside' world is necessary, and certainly across borders, but also when several groups need to report to a central facility, then the needs arises to create a reference list with words, descriptions and, in order to process it in IT-systems, attach meaningless unique codes to these words.
In other words a Reference Terminology is necessary for safe semantic interoperability for nation wide reporting and cross-border exchange of data.

Description of coding systems used in epSOS[53]

epSOS has investigated many coding systems used in Europe and made choices. Choices that had to fulfill a list of criteria.

11.6.c. SNOMED-CT[54]

SNOMED is a Systematized Nomenclature of Medicine-Clinical Terms.
SNOMED Clinical Terms (SNOMED CT) is the most comprehensive, multilingual clinical healthcare terminology in the world.

[53] This chapter is largely based on:
epSOS D3.5.2 Appendix D epSOS Master Value Set Catalogue v0.0.3 20100215.doc

[54] http://www.ihtsdo.org/snomed-ct/

SNOMED CT contributes to the improvement of patient care by underpinning the development of Electronic Health Records that record clinical information in ways that enable meaning-based retrieval. This provides effective access to information required for decision support and consistent reporting and analysis. Patients benefit from the use of SNOMED CT because it improves the recording of EHR information and facilitates better communication, leading to improvements in the quality of care.

SNOMED CT is owned, maintained and distributed by the International Health Terminology Standards Development Organisation (IHTSDO). The IHTSDO is a not-for-profit association which is owned and governed by its national Members. In January 2012 eighteen countries were Members of IHTSDO, more countries are joining every year.
The International Health Terminology Standards Development Organisation is an international not-for-profit organization based in Denmark. IHTSDO owns and administers the rights to SNOMED CT and related terminology standards.

The purpose of IHTSDO is to develop, maintain, promote and enable the uptake and correct use of its terminology products in health systems, services and products around the world, and undertake any or all activities incidental and conducive to achieving the purpose of the Association for the benefits of the members.

The IHTSDO seeks to improve the health of humankind by fostering the development and use of suitable standardized clinical terminologies, notably SNOMED CT, in order to support safe, accurate, and effective exchange of clinical and related health information. The focus is on enabling the implementation of semantically accurate health records that are interoperable. Support to Association Members and Licensees is provided on a global basis allowing the pooling of resources to achieve shared benefits

IHTSDO has more than 25 members.

11.6.d. ICD.X

- ICD: International Statistical Classification of Diseases and Related Health Problems
- Copyright & Issuer: WHO
- Languages & Localization: ICD-10 is available in the six official languages of WHO (Arabic, Chinese, English, French, Russian and Spanish) as well as in 36 other languages.
- ICD-9 CM coding system and the application guidelines are available in Italian.
- Localizations have been made mainly for reimbursement purposes:
- Various manifestations: ICD-10-GM Version 2009 [1.2.276.0.76.5.356] icd10gm2009
- ICD-10-CM [2.16.840.1.113883.6.90]
- ICD-9-CM [2.16.840.1.113883.6.2]
- The Official Updates to the published volumes of ICD-10 are available as annual lists of changes and new versions with slight differences.
- Fields of application: The ICD is the international standard diagnostic classification of diseases (signs, symptoms, conditions) for all general epidemiological, health management and statistics (death) and clinical use (health records). These include monitoring of the incidence and prevalence of diseases and other health problems.
- Number of entries: ca. 13.000 classes
- Structure: 22 chapters, its nodes denote classes of diseases and related problems. ICD classes are arranged into up to five levels. There is one terminal class for each entity. Attributes of ICD are inclusions and exclusions and glossary-like text.

11.6.e. LOINC

- LOINC: Logical Observation Identifiers names and codes.
- Copyright & Issuer: Regenstrief Institute, Indiana
- Languages & Localization: English, Spanish, Chinese, German (Users guide), Estonia Italian
- Fields of application: The scope of the LOINC effort includes laboratory and other clir ical observations. The laboratory portion of the LOINC database contains the usual cat- egories of chemistry, hematology, serology, microbiology, toxicology; as well as categ ries for drugs and the cell counts and antibiotic susceptibilities. The clinical portion of the LOINC database includes entries for vital signs, hemodynamics, intake/output, elec trocardiogram (EKG), obstetric ultrasound, cardiac echo, urologic imaging, gastroen- doscopic procedures, pulmonary ventilator management, selected survey instruments (e.g. Glascow Coma Score, PHQ-9 depression scale, CMS-required patient assessment instruments), other clinical observations and document sections.
- Number of entries: 53,344
- Structure: Each LOINC record corresponds to a single test result or panel. The record includes fields for specifying:
    - Component (analyte) - e.g., potassium, hemoglobin, hepatitis C antigen.
    - Property measured - e.g., a mass concentration, enzyme activity (catalytic rate).
    - Timing - i.e., whether the measurement is an observation at a moment of time, or a observation integrated over an extended duration of time - e.g., 24-hour urine.
    - The type of sample - e.g., urine, blood.
    - The type of scale - e.g., whether the measurement is quantitative (a true measure- ment) ordinal (a ranked set of options), nominal (e.g., E. coli; Staphylococcus au- reus), or narrative (e.g., dictation results from x-rays).
    - Where relevant, the method used to produce the result or other observation

LOINC is also used for the CDA (Clinical Document Architecture) sections

11.6.f. HL7

[To be added]

11.6.g. UCUM

- UCUM: Unified Code for Units of Measure.[Datatype PQ]
- Copyright & Issuer: Regenstrief Institute, Indiana
- Languages & Localisation: English
- Fields of application: UCUM is a system of codes for unambiguously representing measurement units to both humans and machines.
- Number of entries: 556
- Structure: Each unit is defined relative to a system of base units by a numeric factor and a vector of exponents by which the base units contribute to the unit to be defined. Although we can reflect all the meaning of units covered by dimensional analysis with this vector notation, UCUM does not show vectors. Proposed definition from Regenstrief: The Unified Code for Units of Measure is a code system intended to include all units of measures being contemporarily used in international science, engineering, and business. The purpose is to facilitate unambiguous electronic communication of quantities together with their units.

11.6.h. EDQM

EDQM: Standard Terms of European Directorate of Quality in Medicine.

- Copyright & Issuer: EDQM
- Languages & Localization: Albanian, Bulgarian, Chinese, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, German, Greek, Hungarian, Icelandic, Italian, Latvian, Lithuanian, Macedonian, Maltese, Norwegian, Polish, Portuguese, Romanian, Russian, Serbian, Slovak, Slovenian, Spanish, Swedish and Turkish.
- Fields of application: The List of Standard Terms covers dosage forms, routes of administration and containers used for medicines for human and veterinary use.
- Number of entries: ~ 450
- Structure: Part 1: Pharmaceutical dosage forms and short terms; Part 2: Routes of administration; Part 3: Containers.

11.6.i. ATC

- ATC: Anatomical Therapeutic Chemical (ATC) classification system
- Copyright & Issuer: WHO Collaborating Centre for Drug Statistics Methodology, Norwegian Institute of Public Health
- Languages & Localization: English, Spanish, German, Italian
- Various manifestations: ATC-WHO, ATC-GM (Germany), ATC-WIDO (Germany), ATC Vet, ATC Herbal
- International non-proprietary names (INN) are used. If INN names are not assigned, USAN (United States Adopted Name) or BAN (British Approved Name) names are usually chosen.
- Fields of application: In the ATC-classification pharmacological substances are divided into different groups according to the organ or organ system which they affect and their chemical, pharmacological and therapeutic properties.
  A defined daily dose is assigned to each active substance. Defined daily doses (DDD) are the assumed average daily maintenance dose for the main indication of each substance in adults. The ATC-Classification with defined daily doses serves as an easing of comparisons between drugs and guarantees a standardized reference for the specification of daily treatment expenses. The purpose of the ATC/DDD system is to serve as a tool for drug utilization research in order to improve quality of drug use.
- Number of entries: 4067 Codes (2006)
- Structure: In the Anatomical Therapeutic Chemical (ATC) classification system, the drugs are divided into different groups according to the organ or system on which they act and their chemical, pharmacological and therapeutic properties. Drugs are classified in groups at five different levels. The drugs are divided into fourteen main groups (1st level), with one pharmacological/therapeutic subgroup (2nd level). The 3rd and 4th levels are chemical/pharmacological/therapeutic subgroups and the 5th level is the chemical substance. The 2nd, 3rd and 4th levels are often used to identify pharmacological subgroups when that is considered more appropriate than therapeutic or chemical subgroups.

  ATC assures the unified coding of active components registered. This feature, coupled with strength (dosage) and pharmaceutical form) assures the possibility to transfer the full information on a medication from Country to Country, regardless the brand name of the medication. A medicinal product can be given more than one ATC code if it is available in two or more strengths or formulations with clearly different therapeutic uses. The existence of multiple codes reduces the risk of mistakes in the specification of strength and prescription / dispensation of different medication for the two pathologies.

11.6.j. Coding systems: Choices made by epSOS[55]

epSOS has investigated several potential coding systems.
The next table shows choices made for coding system that are used at present in epSOS. These coding systems have been described in the previous chapter.

SNOMED codes are used in place where no license-free alternative was present.

epSoS CDA R2	
Field	**Terminology Binding**
Field Labels	LOINC
Problem list	ICD 10
Medication list	ATC + EDQM + UCUM
Allergies	Allergen: SNOMED Active ingredient, medicaments: ATC
Surgical procedures	SNOMED
Medical devices	SNOMED
Country and languages	ISO 639
Professional role	ISCO
Vaccinations	tbd
SNOMED CT not licensed in all countries.	

11.7. SAM: SNOMED coding systems as used

The epSOS dataset was transformed into an ISO 13606 based expression. As much as possible SNOMED code were to be used.

[55] This chapter is largely based on:
epSOS D3.5.2 Appendix D epSOS Master Value Set Catalogue v0.0.3 20100215.doc

The agreed process was to annotate the epSOS SAM with relevant SNOMED codes and the NHS-UK Coding centre would validate these codes.
During that annotation process it appeared that SNOMED was missing many primitive codes. E.g. there was a code for the concept 'Patient Name' but no code for the concept 'Name'.
At the same time in the EU-SemanticHealthNet project experts discussed the same topic. One of the options that will be considered is to ask IHTSDO to add the needed primitive SNOMED codes. Primitive SNOMED codes that can be used to annotate the SIAMM generic pattern.

ERS tried to obtain SNOMED-CT codes for the HSE version of the epSOS dataset and in particular the nodes in the 13606 Template and archetypes that need codes.
In discussions with the UK Terminology Coding Centre and with experts from IHTSDO is became clear that only codes for clinical concepts were within the scope but not the general (primitive) codes. At present a discussion has been started with IHTSDO about this problem of missing primitive codes.

12. Appendix: Glossary

Term	Description	Comment
CDA	Clinical Document Architecture	HL7 and ISO specification of documents with EHR-data contents
CEN	European Standard Development organisation (SDO)	
epSOS	European Patient Smart Open Services	a project by EU-Member states for cross border exchange of the Patient summary and Prescriptions.
Health record extract	part or a health record grouped for the purpose of communication	
HL7	SDO from the US mainly producing message standards (HL7v2, HL7v3 and HL7v3 CDA R1 and R2)	
IA-RM	Information Architecture Reference Model	a term coined by HSE to describe how artefacts are produced that express data needs and can be used in procurement
ICD-x	International Statistical Classification of Diseases and Related Health Problems	A classification published and maintained by the WHO
IHE	Integrating the Healthcare Enterprise	an organisation needed to make message profiles and to test implementations of messages
IHTSDO	SDO that produces and markets SNOMED-CT	SNOMED-CT is the only international reference terminology
INFOstructure	all services needed for semantic interoperability	
ISO	International Standard Development organisation (SDO)	
LOINC	Logical Observation Identifiers Names and Codes	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.
Patient Summary	electronic health record extract that provides an electronic patient health dataset applicable both for unexpected, as well as expected, healthcare contact	electronic health record extract that provides an electronic patient health dataset applicable both for unexpected, as well as expected, healthcare contact
Reference Model	a formal model used to generate outputs	
SAM	Semantic Area Model	A formal specification of a collection of data points for a specific purpose and users (both healthcare providers, authorities and IT-specialists

Term	Description	Comment
UCUM	Unified Code for Units of Measure	a standard produced by the Regenstrief Institute, Inc. and the UCUM Organization, Indianapolis, IN
WHO	World Health Organisation	

Appendix: Questionnaire