E-Health in Belgium and in the Netherlands
F.H. Roger France et al. (Eds.)
IOS Press, 2002

Towards Developing a Coherent Healthcare Information Infrastructure

A.C.M. DUMAY & G. FRERIKS
TNO Prevention and Health, Division Technology in Health Care
Zernikedreef 9, 2333 CK LEIDEN, The Netherlands

Abstract. A standard is a set of agreements between all parties involved. Delivering healthcare is a matter of co-operation: healthcare can only be delivered in a responsible fashion when parties involved comply to standards. This becomes evident when a healthcare information infrastructure is under development. The authors deliver a comprehensive introduction to standards which apply to the Electronic Healthcare Record (EHR), describe work in progress of relevant standardisation committees and argue about future work of such committees in order to pursue a coherent healthcare information infrastructure.

Keywords. Healthcare Information Infrastructure Architecture, Electronic Healthcare Record, Standards, Archetypes.

1. Introduction

The International Organisation for Standardisation ISO defines standards as [1]:
- documented agreements
- containing technical specifications or other precise criteria;
- to be used consistently as rules, guidelines, or definitions of characteristics;
- to ensure that materials, products, processes and services are fit for their purpose.

The measure of accuracy, precision and consistency of standards and their application has an enormous impact on (social-) technical systems. Exchanging information is important and necessary in today's healthcare practice, where many medical professions and supporting services come together. Information exchange is supported by technology. For many years this was pen and paper. The arrival of computer systems in the medical profession created an opportunity to document healthcare information in an electronic fashion. With the use of computer systems did the wish arise to register, store, retrieve, present, process and transport healthcare information within a medical organisation such as the hospital. Further professional specialisation and differentiation of tasks and roles urged healthcare providers to seek co-operation with healthcare providers outside their own organisational unit. The upscale to regional co-operation and, recently, co-operation at national level forced the healthcare community to take appropriate measures to manage the exchange of information within the great diversity of organisations, installed computer systems, connections and software applications. The need of mutual agreements between parties involved is also evident: the need of standards to support developing and implementing a coherent information infrastructure.

The authors deliver a comprehensive introduction to standards which apply to the Electronic Healthcare Record (EHR), describe work in progress of relevant standardisation committees and argue about future work of such committees in order to

pursue a coherent healthcare information infrastructure. Section 2 deals with types of information exchange, the characteristics and features of information exchange and context. Section 3 of this paper is devoted to technology and the evolution of standards that apply to the Electronic Healthcare Record (EHR) and highlights work in progress of relevant standardisation committees. The discussion can be found in section 4.

2. Information exchange in healthcare

This section deals with types of information exchange: messages, documents, objects; the characteristics and features of information exchange in the frame of the ISO/OSI model; and the context of information exchange.

The ISO/OSI model [2] describes system interconnections in 7 layers. The seventh layer is referred to as the application layer. To complete system interconnections into system interoperability an additional layer should be incorporated: the semantics layer, to agree on classification and terminology. At each layer a number of standardisation committees deliver effort to increasing precision, accuracy and consistency. They plough through details in order to facilitate open healthcare system interconnection.

Communicating parties exchange information. Communication is possible when at least a minimum set of agreements or conditions are satisfied. Table 1 compares agreements for the use of paper-based and computer-based communications.

Table 1. Conditions to comply in case of paper-based and computer-based communication.

Paper-based communication	Computer-based communication
The minimal content.	The minimal data sets defined by scientific communities.
The coding system(s), terminology and dictionary.	The coding and classification system(s), e.g. ICPC-2, SNOMED-CT, ICD-10.
The form of the information (memo, report, letter of discharge, etc.).	The form of the information, usually the combination of the minimal content, structure and form.
The syntax of the language used.	A choice of syntaxes to register information, e.g. CEN/TC251 EN 13606, HL7 v3 RIM and CDA, OMG COAS, DICOM Structured Reporting.
The way of writing (ink, pencil, colour).	The way of writing, e.g. Edifact/X12, XML, ASN.1 and also ASCII and UNICODE. Agreements borrowed from outside the healthcare domain.
The paper format and size.	The format of files, email, etc. Also imported into the healthcare domain.
The envelope and address.	The envelope information like address, sender, date and time of transmission, etc. The ebXML consortium (i.e., UN/Cefact and OASIS) defined ways of electronically enveloping electronic information.
The shipping of the envelope (including its content) and services for routing, tracking & tracing, delivering at the specified address at the specified time, confirmation of arrival, registration, postal code books, address books, etc.	The shipping of the envelope with content as laid down in standards by OMG, CEN/TC251 and HL7.
Agreement on who is authorised and why what to do with the content.	Agreement on who is authorised and why what to do with the content, defined by e.g. CEN/TC251, ISO/TC215, W3C, IETF and ebXML.

Defining minimal content, coding system(s), terminology and dictionaries is the area of healthcare professionals and their scientific boards. Supporting technical standards are available from CEN/TC251 [3] and HL7 [4]. The syntaxes and information carrier formats are also, but not exclusively, in the area of healthcare professionals. CEN/TC251 and HL7 but also OMG [5] and DICOM [6] are active in this field, yielding quite different standards. Formats and sizes are found at several layers of the ISO/OSI model are not specific for healthcare and transmitted in various forms, like printed paper, faxed, by X400 protocol, e-mail and digital objects not exclusive to healthcare. CEN/TC251 and HL7 will adopt envelope standards delivered by ebXML [7], UN/Edifact [8] and OASIS [9] which are also not exclusive to healthcare. Shipping or transporting information is in the area of the healthcare information infrastructure. This is the area where physical and virtual space meet. ISO/OSI describes the components involved in transporting information. The healthcare information infrastructure is completed with standards on information security (including the Public Key Infrastructure), tracking & tracing of information and retrieval of addresses. Non-healthcare and healthcare specific standards are combined. Agreements on authorisation are the area of regulatory reforms focussing on defining and allocating responsibilities.

3. Standardisation work in progress

This section deals with the evolution of technology relevant to the Electronic Healthcare Record and focuses on those technologies on which standardisation committees have a different view.

3.1. The Reference Information Model

Common practise in deriving at standard messages was starting with a minimal data set (WHAT to communicate), agreeing on communication scenario's (HOW to communicate) and composing these into a Domain Information Model. The domain model is then used to set up a hierarchical list of messages which were implemented in Edifact format. This approach was followed by many user communities and yielded many domain information models where (partly) the same data were communicated and ended in conflicting models and deviation from accepted standards.

In 1996, HL7 decided to compose a Reference Information Model (RIM) from all available domain models. This new and coherent meta-model is the basis for specific new domain models: the RIM ensures consistency among standards.

3.2. Documents

Standard messages are characterised by its content, the minimal data set. This applies to all messages by CEN/TC251 and HL7. As a result, many standard messages came about: a laboratory message, a referral message, a discharge message, an obituary message, etc. The messages can be regarded to support updating the status of two (or more) communicating systems.

The eXtensible Mark-up Language (XML) triggered standardisation committees to take the concept of documents into consideration. Documents are artefacts to communicate information among people rather than databases. Documents are to be seen as electronic letters, used in the correspondence between healthcare professionals, with the ability to be stored in computer systems. Documents consist not only of content but also of instructions on how to present the content and the signature of the healthcare professional who is responsible for the content. By developing document standards there is no longer a need for a detailed description of the content, as practised for message standards. A document can be a laboratory message, a series of messages or a full healthcare record. The document standard has become a model to register, store, retrieve, present, process and transport healthcare information in general.

CEN/TC251 developed the first document standard in the frame of ENV 13606 [10] and ENV 13607 [11] in order to communicate documents by XML, the eXtensible Mark-up Language.

3.3. Extensible mark-up language

Mark-up is labelling text with instructions for the presentation or the meaning of that text. During the 1970's mark-up was developed and effort resulted in the Standardised General Mark-up Language (SGML) which became an ISO standard [12]. The World Wide Web Consortium (W3C) [13] started to simplify SGML in 1996 and yielded the eXtensible Mark-up Language XML.

XML as an exchange format made it possible to develop message standards and document standards which are flexible and easy to use compared to usual exchange formats as Edifact and X12. The latter standards were hampered by the fact that the position of data in the message or document determined its meaning.

Both CEN/TC251 and HL7 decided in 1996 to prefer XML as information exchange format.

3.4. Archetypes and standard models

Australian researchers introduced recently the concept of archetypes [14]. An archetype can best be described as documents in the above meaning. Archetypes are Clinical Concept Models. The archetype defines in a uniform and coherent way the information related to a concept to be stored in databases. Communities of healthcare professionals decide on the minimal data sets in concepts it needs to communicate including presentation rules. Examples are: blood pressure value (systolic, diastolic), apgar score, referral message, liver laboratory test, etc. They are supported by predefined templates which are part of the archetype standard. Building archetypes is supported by a so-called archetype editor which keeps all the rules of the standard, the archetype meta model, and information technology artefacts like data types and facilities for persistence, the so-called kernel. CEN/TC251 has adopted these two model of the Good Electronic Healthcare Record (GEHR).

Archetypes create the major advantage that minimal data sets remain the responsibility of communities of healthcare professionals or their scientific boards rather than ICT

suppliers. Suppliers should focus on ICT functionality based on the archetypes, the communities become the registries of archetypes. Another major advantage is the support of legacy. Archetypes (or data sets) which become obsolete as a result of organisational change, substitution of test methods in the laboratory environment and so on, can exist next to each other. This means that the way of storing legacy concept models and new concept models does not affect existing databases.

3.5. Communication forms

There are basically three forms of communication possible.
First, there is visual integration. Information is presented onto a computer screen without the option to register, retrieve, process, store or transport the information. This type of communication is applied in, e.g., laboratory applications. Its analogy in the world outside the healthcare domain is the Fax.

Secondly, there is the exchange of messages and documents. Both are artefacts of asynchronous communication, i.e., the sender and the reader do not need to be involved in the communication process at the same time. Messages and documents can be registered, presented (documents can even be presented as letters), processed or transmitted. Message and document standards assume that the sending and receiving party share only a uniform model with which the message or document has been build. Such a model, as delivered by HL7 version 2, the forthcoming HL7 version 3 and the 'old' CEN/TC251 standards, is not meant to support storage in computer systems. This type of information exchange is called loose coupling.

Thirdly, there is the exchange of objects that takes place in a synchronous fashion: sender and receiver are connected at the same time to exchange information. Recently, CEN/TC251 decided to fully support storage and exchange of medical information through objects. The requirement to store medical information in a standard format means that the model to define objects and messages & documents cannot be the same. Only when systems make use of the same domain model for the EHR derived from a standard is object exchange possible. It is always possible to convert objects into messages and documents and use them in asynchronous communication. The main advantage of information exchange in objects is that systems can be coupled tightly and provide inherent stronger information security.

In the context of information exchange in objects is the feature of early and late binding. Early binding stands for freezing the format of the communication artefact during compilation of the source code. Messages are examples of early binding. Consequently, change of the messages requires compilation of the source code and installing a new software release. Late binding allows users to download a specific structure at the moment of sending or receiving a message. Systems which exchange information by XML documents based on CEN/TC251 and the forthcoming HL7 version 3 can carry this feature.

3.6. Co-operation and harmonisation

From the early 1990's, CEN/TC251 started co-operating with other standardisation committees like those from HL7 and DICOM. CEN developed parts of the DICOM

standard for DICOM: the DICOM standard is now accepted as a European standard. CEN/TC251 and HL7 deploy a similar method to derive at messages and each adopted the best of each other's method. CEN/TC251 borrowed the interesting HL7 version 3 Reference Information Model and up to now derived two standards with the HL7 version 3 RIM as meta-model. Experience learned that today's version of the RIM is complete so that all requirements of CEN/TC251 message standard can be complied with. The basic HL7 data types were adopted by CEN/TC251, a few attributes were added. This harmonisation process between CEN/TC251 and HL7 was possible as a result of a Memorandum of Understanding between the two organisations.

3.7. Standardisation products

Standardisation organisations like CEN/TC251 and HL7 delivered message standards for over the last 10 years or so. The message standards contained at least the minimal data set and its structure. The standards resulted from a consensus process, involving many people from many countries, and consequently featured many options. Although the standards were rigid, the features allowed people to create regional or national profiles to the standard.

Recently, a component-based approach is followed. The components serve as building blocks for users to create local message standards. Since CEN/TC251 and HL7 started applying a meta model, like the Reference Information Model, to derive at standards of various kind it is expected that the internal meta models will be subject to standardisation as well. Or, as is the case with the General Purpose Information Components (GPIC's) of CEN/TC251, these building blocks derived from the meta model will become the product of standardisation. (The CEN/TC251 GPIC document is offered to ISO/TC215 for co-development.)

4. Discussion and conclusions

Most standards available date from the period where healthcare information was exchanged in a clear context. I.e., between professionals within the same organisational unit or at least in a clear organisational context. Professionals were supported by stand-alone computers, most of the time for billing purposes only. Over time, the healthcare paradigm shifted towards specialisation and co-operation and information and communication became an issue. And with it communication and information technology. Professionals needed technology to register, store, retrieve, present, process and transport healthcare information. First within limited organisational confines, today at regional and national, and even cross-border, levels. Different healthcare professions meet in a variety of organisational constraints, using different generations of computers systems, computer control software and application software. Today, not only efficiency is at stake but patient safety and health professional legal security as well. There is an urgent need to bring standards, methods and organisational constraints together in a new architecture of the healthcare information infrastructure.

A coherent healthcare information infrastructure needs two types of standards: a system architectural standard and a standard for generic services. The CEN/TC251 health information systems architecture (HISA) is a standard of the first type. The OMG standards are of the second type. They define together the healthcare information

infrastructure. CEN/TC251 is unique in the sense that it is the only standard to determine the terms and notions applicable, and is the only standard to handle information as objects. The OMG object broker technical architecture proved to support scalability of infrastructure. Commercial of the shelf and open source products are available. Still, a lot work need to be done to reach coherency in this framework:

- stimulate and facilitate development and application of standards for object handling in healthcare. CEN/TC251 is such a standard which should be leveraged to ISO/TC251 world standard;
- support and facilitate message handling in the object broker technical infrastructure. The infrastructure must support applications exchanging information by messages (HL7 or Edifact). Interfacing of applications should be reached by communication servers;
- revise and facilitate the CEN/TC251 healthcare information infrastructure architecture (HISA) into a European norm and an ISO/TC215 world standard. A taskforce within CEN/TC251 has picked up this issue;
- encourage and facilitate the use of XML over Edifact. XML is world wide accepted as the successor of Edifact;
- stimulate and facilitate the use of HL7 messages. HL7 messages standards combine well with the CEN/TC251 electronic health record. Further harmonisation between CEN/TC251 and HL7 must be encouraged and facilitated;
- prepare and facilitate professional healthcare communities to decide on the minimal data sets to communicate including presentation rules. The archetype editor is publicly available for this purpose. Consequently, ICT suppliers need not modify databases and health professionals need no longer to install new software revision.

References

[1] http://www.iso.org
[2] ISO/OSI 7498, http://www.iso.org
[3] http://www.centc251.org
[4] http://www.hl7.org
[5] http://www.omg.org
[6] http://medical.nema.org
[7] http://www.ebXML.org
[8] http://www.edifrance.org/edifact-wg/
[9] http://www.oasis-open.org
[10] ENV 13606-1/4:2000 Health informatics - Electronic healthcare record communication, http://www.centc251.org/FinWork/greensheetpwd.htm
[11] CEN/TC251 ENV 13607, http://www.centc251.org/FinWork/greensheetpwd.htm
[12] http://www.w3.org/MarkUp/SGML/
[13] http://www.w3.org
[14] http://www.gehr.org