

Survey of Electronic Health Record Standards

Montserrat Batet Sanromà

Aida Valls Mateu

Karina Gibert Oliveras

Research Group on Artificial Intelligence (BANZAI)

Document from the Research Project K4CARE (IST-2004-026968 K4CARE)

Summary: *The aim of this report is to provide an overview of Electronic Health Records (EHR). The increasing importance of EHRs is due to the fact that they must provide a mechanism to unify the large quantity of information about the health care of the patients' diagnosis and treatment. The integration of this information would make possible that the patient care become more effective and efficient. In this survey we present the different EHR standards that are being developed, however, the explanation is focused mainly on the three standards that are taking the leadership: GEHR/openEHR standard, CEN EN 13606 EHRcom standard and HL7 standard.*

1. Introduction

The Electronic Health Record (EHR), also called Electronic Healthcare Record, is an important component in health informatics. It is defined as “digitally stored health care information about an individual’s lifetime with the purpose of supporting continuity of care, education and research, and enduring confidentiality at all times” [Iakovidis, 1998]. An EHR includes all the information derived from the care of a patient, such as observations, laboratory tests, diagnostic imaging, reports, treatments, therapies, drugs administrated, patient identifying information, legal permissions, and allergies.

A more formal definition is given in [ISO/DTR 20514]: “The *basic-generic* definition for the EHR is a repository of information regarding the health status of a subject of care, in computer processable form”. Where, the term *subject of care* is used synonymously with *patient*. It is accepted that the most important characteristic of an EHR has to be the ability to share information between different authorized users. Thus, this requires interoperability of information inside the EHR and interoperability of EHR systems which exchange and share this information. Following this idea, the basic-generic EHR is divided in two specializations, sharable and non-sharable EHRs as illustrated in Figure 1 One of the sharable EHR considered is the "Integrated Care EHR" (ICEHR) as defined in [ISO/DTR 20514]: “The Integrated Care EHR is defined as a repository of information regarding the health of a subject of care in computer processable form, stored and transmitted securely, and accessible by multiple authorized users. It has a 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”.

The ICEHR is based on a standardised or commonly agreed logical information model which supports semantic interoperability. The *openEHR* Reference Model and the CEN 13606 Reference Model are examples of models which fit this definition.

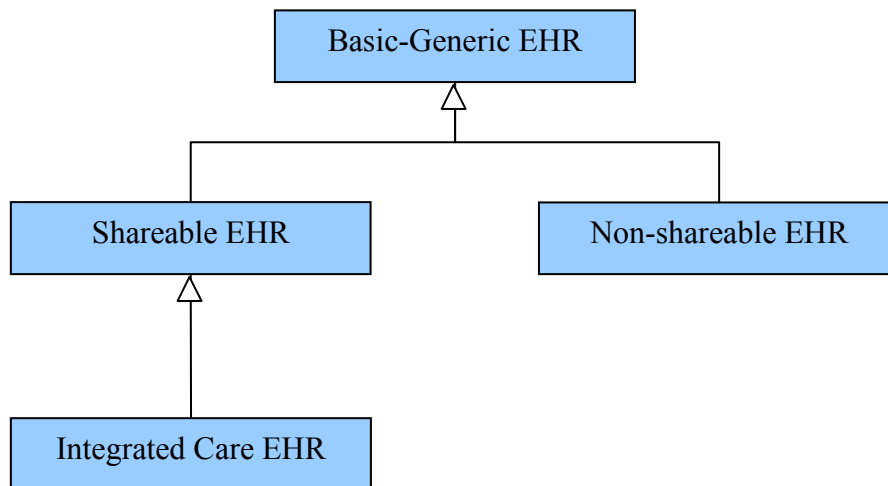


Figure 1 Specialisation of the Basic-Generic EHR

According to [Cohen, 2001], the standards must consider the following aspects:

- healthcare messages exchange
- EHR object model
- healthcare terminology/ vocabulary

Additionally an EHR can provide a proper security mechanism. For example, authentication or encryption.

There are two main organizations that define standards:

- HL7 Health Level 7(www.HL7.org) USA
- CEN TC 251 (www.centc251.org) Europe

There are many standards under development, the ones that are leading the research are the CEN EN 13606 EHRcom and HL7 standards, which come from the GEHR/openEHR standard. To have more widely vision of EHR standards, this report presents also other standards and compares them. This document is organized as follows: in section 2 there are the description of GEHR/openEHR standards, in section 3 and 4, they are describes EHRcom and HL7 standards, respectively. In section 5 there is a brief summary of the standard called Integrating the Healthcare Enterprise. In section 6 the Digital Imaging and Communications in Medicine is introduced. Finally, in section 7 we compare some aspects of the standards and present some conclusions.

2. GEHR/openEHR

The GEHR/openEHR standard project began in January 1992, as a European Union initiative called Good European Health Report. The GEHR Project finished at the end of 1994. Later Australia enters into the project and the initiative was called Good Electronic Health Record. The renewed Australian focus on developing the results of the original GEHR project has had a considerable influence in the ISO, HL7 and CEN.

In 1999, it was established an open source foundation to take forward harmonisation in the field, from patient and clinical perspectives. The name *openEHR* was adopted. Actually, the project is maintained by the openEHR foundation.

The *openEHR* Foundation is an international, on-line community whose aim is to promote and facilitate progress towards electronic healthcare records of high quality, to support the needs of patients and clinicians everywhere. The *openEHR* Foundation [openEHR] was created to enable the development of open specifications, software and knowledge resources for health information systems, in particular electronic health record (EHR) systems. It publishes all its specifications, and builds reference implementations of them, as open source software.

Some of the *openEHR* aims are:

- promote and publish the formal specification of requirements for representing and communicating electronic health record information, EHR information architectures, models and data dictionaries;
- manage the sequential validation of the EHR architectures;
- and maintain open source "reference" implementations.

The *openEHR*'s information model is the **archetype**. An archetype is a re-usable, formal model of a domain concept. As defined in [Beale and Heard, 2005] an *archetype* is "a computable expression of a domain content model in the form of structured constraint statements, based on some reference model". *openEHR* archetypes are based on the *openEHR* reference model. This model describes de health record itself. It is composed by packages. Each package defines one *openEHR* specification document, an information model. For example, the EHR information model [Beale et al, 2006] is organized in folders and compositions. Compositions are a boarder concept than documents, but include documents. Examples of Compositions are a progress note, or a laboratory report. The Composition is the EHR's top level data container. Folders can be used to classify Compositions in a hierarchy. The package EHR contains the top level structure of the EHR. The package Content contains the Entry package (which is used to record clinical statements) and the Navigation package (which contains sections that provides a structure to the record similar to headings). In Figure 2 there is a graphical representation of the package structure of the *openEHR* EHR information model.

In general, archetypes are defined for wide re-use, however, they can be specialized to include local particularities. They can accommodate any number of natural languages and terminologies. In healthcare, an archetype can model concepts such weight measurement, blood pressure, microbiology results, prescription, diagnosis and so on.

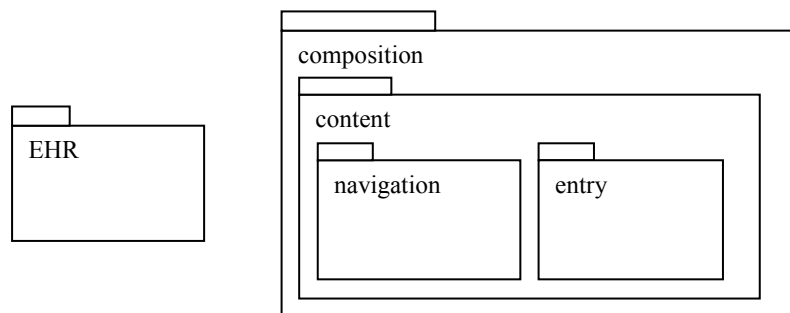


Figure 2. Package structure of EHR information model

The formal language for expressing archetypes is ADL (Archetype Definition Language) [Beale and Heard, 2006]. ADL is a knowledge description language. ADL uses three other syntaxes to describe constraints on data which are instances of some information model, cADL (constraint from ADL), dADL (data definition from of ADL), and a version of first-order predicate logic (FOPL). In Figure 3 there is the ADL Archetype Structure. Sections in brackets are optional.

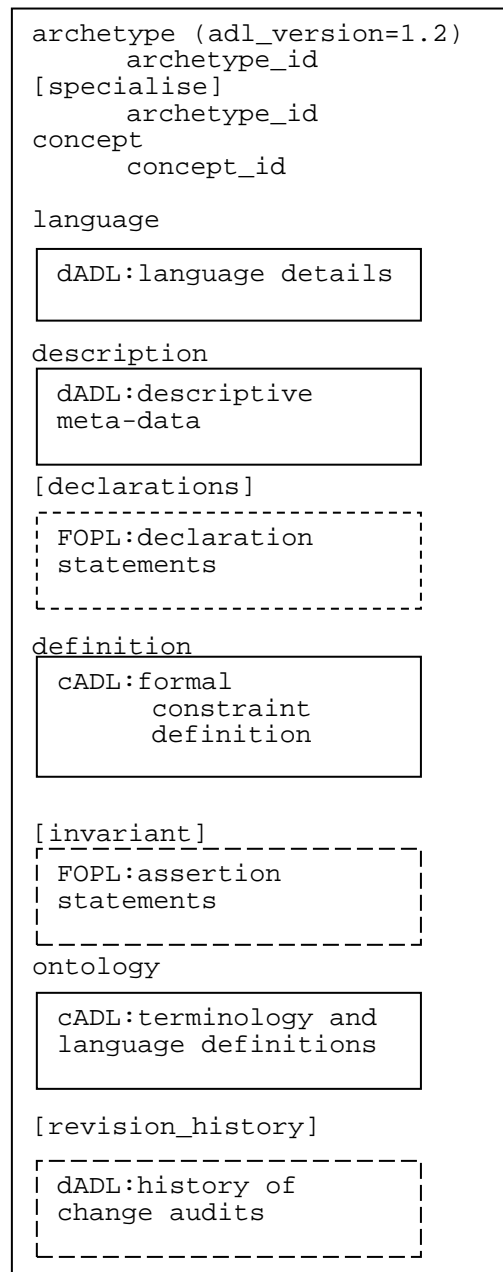


Figure 3. ADL Archetype Structure

An archetype is divided into three main parts:

- **Descriptive part** contains a unique identifier, a machine-readable code describing the clinical concept modelled by the archetype and various metadata

such as version, the author's name or the organization, the language used or the purpose.

- **Definition part** is the main part of an archetype. Describes the architecture, content or restrictions of the archetype.
- **Ontology part** defines the vocabulary. It may contain language translations of code meanings of codes used within the archetypes to external vocabularies such as SNOMED[SNOMED] or LOINC [LOINC], ICD (International Classification of Diseases) [WHO, 2006], ICPC (International Classification of Primary Care) [WHO, 2006].

The following example shows an archetype definition expressed in ADL describing microbiology laboratory observations.

```

archetype (adl_version=1.4)
  openEHR-EHR-OBSERVATION.microbiology.v1
concept
  [at0000] -- Microbiology laboratory observations

description
  original_author = <
    ["name"] = <"Sam Heard">
    ["organisation"] = <"Ocean Informatics">
    ["date"] = <"2004-05-18">
    ["email"] = <"sam.heard@oceaninformatics.biz">
  >
  details = <
    ["en"] = <
      language = <"en">
      purpose = <"microbiology observations for a single specimen. Multiple specimens should use
multiple entries.">
      use = <"For reporting the objective findings of a sample submitted for microbiological
assessment">
      keywords = <"microbiology", "culture", "organism">
      misuse = <"This does not form the complete microbiology report, a composition - which may
require evaluations such as diagnosis, recommendations and referrals.">
    >
  >
  lifecycle_state = <"AuthorDraft">

definition
OBSERVATION[at0000] ∈ { -- Microbiology laboratory observations
  data ∈ {
    HISTORY[at0002] ∈ { -- history
      events cardinality ∈ {1..*; unordered} ∈ {
        EVENT[at0003] occurrences ∈ {1..*} ∈ { -- specimen collection time
          data ∈ {
            ITEM_TREE[at0001] ∈ { -- structure
              items cardinality ∈ {0..1; ordered} ∈ {
                ELEMENT[at0004] occurrences ∈ {0..1} ∈ { -- Result label
                  name ∈ {
                    CODED_TEXT ∈ {
                      code ∈ {[ac0000]} -- =label, battery name
                    }
                  }
                  value ∈ {
                    TEXT ∈ {*}
                  }
                }
              }
            }
          }
        CLUSTER[at0005] occurrences ∈ {0..1} ∈ { -- Specimen
          items cardinality ∈ {0..1; ordered} ∈ {
            ELEMENT[at0006] occurrences ∈ {0..1} ∈ { -- Sample description
              value ∈ {
                TEXT ∈ {*}
              }
            }
          }
        }
      }
    }
  }
}

```

```

    }
  }
  ELEMENT[at0007] occurrences E {0..1} E { -- Site
    value E {
      TEXT E {*}
    }
  }
  ELEMENT[at0008] occurrences E {0..1} E { -- Sampling method
    value E {
      TEXT E {*}
    }
  }
  ELEMENT[at0036] occurrences E {0..1} E { -- Sterile field
    value E {True, False}
  }
}

...
}
}
}
}
protocol E {
  ITEM_LIST[at0030] E { -- protocol structure
    items cardinality E {0..1; ordered} E {
      ELEMENT[at0031] occurrences E {0..*} E { -- Microscope type
        value E {
          TEXT E {*}
        }
      }
      ELEMENT[at0032] occurrences E {0..1} E { -- Stain
        value E {
          CODED_TEXT E {
            code E {[local::]}
          }
        }
      }
    }
  }
}
...
}
}
}
}

```

ontology

```

primary_language = <"en">
languages_available = <"en", ...>
term_definitions = <
  ["en"] = <
    items = <
      ["at0000"] = <
        description = <"The observations made in the microbiology laboratory based on a single specimen. This archetype can be repeated in one full microbiology report">
        text = <"Microbiology laboratory observations">
      >
    >
  >
  >
  >
constraint_definitions = <
  ["en"] = <
    items = <
      ["ac0000"] = <
        description = <"The name or label given to the result - determined by the laboratory">
        text = <"=label, battery name">
      >
    >
  >

```

```

>
["ac0001"] = <
  description = <"A microscopy field e.g. low power field, high power field etc">
  text = <"Any term that 'is_a' microscopy field">
>
["ac0002"] = <
  description = <"Any micro-organism that will be reported in a laboratory">
  text = <"Any term that 'is_a' micro-organism">
>
["ac0003"] = <
  description = <"Any antimicrobial agent for which sensitivity testing is (or could be)
undertaken">
  text = <"Any term that 'is_a' antimicrobial therapy">
>
["ac0004"] = <
  description = <"A label for the adjustment made in the result">
  text = <"Any valid adjustment label">
>
>
>
>

```

3. CEN EN 13606 EHRcom

CEN ENV/EN 13606 EHRcom is developed by CEN/TC 251, which is the technical committee on Health Informatics of the European Committee for Standardization. Its scope is standardization in the field of Health Information and Communications Technology to achieve compatibility and interoperability between independent systems and to enable modularity [CEN/TC 251].

This standard is based in EHR exchange messages and adopts the archetype methodology of *openEHR*

EHRcom will be a standard consisting of five parts:

- **Reference Model:** defines the hierarchy of generic building blocs of the EHR through a set of classes. Represents the stable characteristics of the EHR entries, how they are aggregated, and the context information required to meet ethical, legal and provenance requirements and would be embedded in a distributed EHR environment as specific messages or interfaces. This model needs a formal method to organize and share the predefined classes of the EHR. This model corresponds conceptually to the EHCR architecture of GEHR, the Synapses SynOM, the information model of ENV 13606-1 and the openEHR Reference Model.
- **Archetype Interchange Specification:** each archetype defines legal combinations of the building-block classes defined in the Reference Model for particular clinical domains, organisations. The archetype model is syntactic equivalent to of the Good Electronic Health Record project, and openEHR. Archetypes are expressed in this ADL and will also be convertible to HL7 RMIMs and CMETs.
- **Reference Archetypes and Term Lists:** includes the vocabularies for attributes, and archetypes to represent HL7 specialised Acts and *openEHR* specialised ENTRYs.

- **Security Features** :defines an interoperable specification for EHR disclosure consent, and an interoperable disclosure log
- **Exchange Models**: this part is under discussion at the moment.

The EHRcom reference model is divided into several class packages [CEN prEN 13606-1, 2006].

- The Extract package, which defines the EHR_EXTRACT root class of the reference model and the EHR data that it contains (see Figure 4).
- Demographics package, which provides a minimal data set to define the various persons, software agents, devices and organisations that are referenced within the EHR_EXTRACT.
- The Terminology package includes the relevant definitions of terms used within the EHR in order to permit the interpretation of the EHR_EXTRACT.
- A set of data type packages defining the representation of attribute and data values for quantities, text, primitive and basic types.
- The Access Control package, which is under development in EN 13606-4. It will define a representation for EHR access policies.
- The Message package; this class is a placeholder for the attributes that will be required to communicate the EHR_EXTRACT to a requesting process via a message or other serialised form. This package includes HL7 Domain Message Information Model (D-MIM) to allow the use of HL7 for communication of EHR extracts.

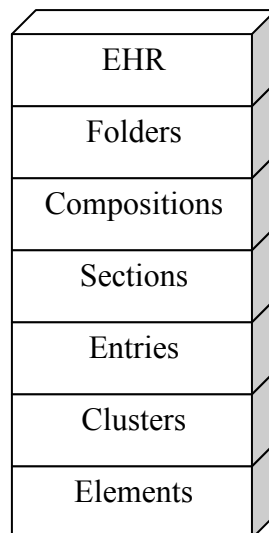


Figure 4. Components of the EHRcom Extract Reference Model

An important part of the standard is the information model adopted to structure the EHR information. In this standard, the data is organised into a hierarchy with seven levels (see Figure 4) [CEN prEN 13606-1, 2006]. The EHR_EXTRACT is the top-level component of the EHR of a single subject of care. Within an EHR the FOLDER is the

highest level in the hierarchy. The FOLDER divide the EHR into compartments representing a disease, a clinical team, an institution or a fixed time period (p.e. Diabetes, Schizophrenia, Ophthalmology, Cardiology, Pius’s Hospital, Episodes 2000-2001). A COMPOSITION is the information of a single clinical encounter or record documentation session (p.e. Laboratory test, Radiology report, Diabetes review). A SECTION is information within a COMPOSITION that belongs under one clinical heading. It reflects the information gathered in a clinical encounter (p.e. Reason for encounter, Past history, Family History, Allergy information, Subjective symptoms, Objective findings, Analysis, Plan, Treatment, Diet, Posture, Abdominal examination, Retinal examination). AN ENTRY (Figure 5) is the information recorded in an EHR as a result of one clinical action, observation, clinical interpretation, or an intention (p.e. a symptom, an observation, one test result, a prescribed drug, an allergy reaction, a diagnosis, a differential diagnosis, a differential white cell count, blood pressure measurement).

ENTRY Blood Pressure Graph Query

CLUSTER: Query Specification

ELEMENT: Query Syntax: <EHR_OQLv1>

ELEMENT: Query String: “Select...

where Cluster.meaning = <Blood Pressure>

and containing.Entry.subject_of_information = <Patient>

and containing.Composition.Clinical_Session.session_time.start

> (now>-365days)”

ELEMENT: Datetime first authored: 20 February 2003

Figure 5. Example of entry in EN 13606

CLUSTER is used to organize nested multi-part data structures such as time series, and to represent the columns of a table (p.e. Audiogram results, electro-encephalogram interpretation, weighted differential diagnoses). ELEMENT contains a single data value (p.e. Systolic blood pressure, heart rate, drug name, symptom, body weight.).

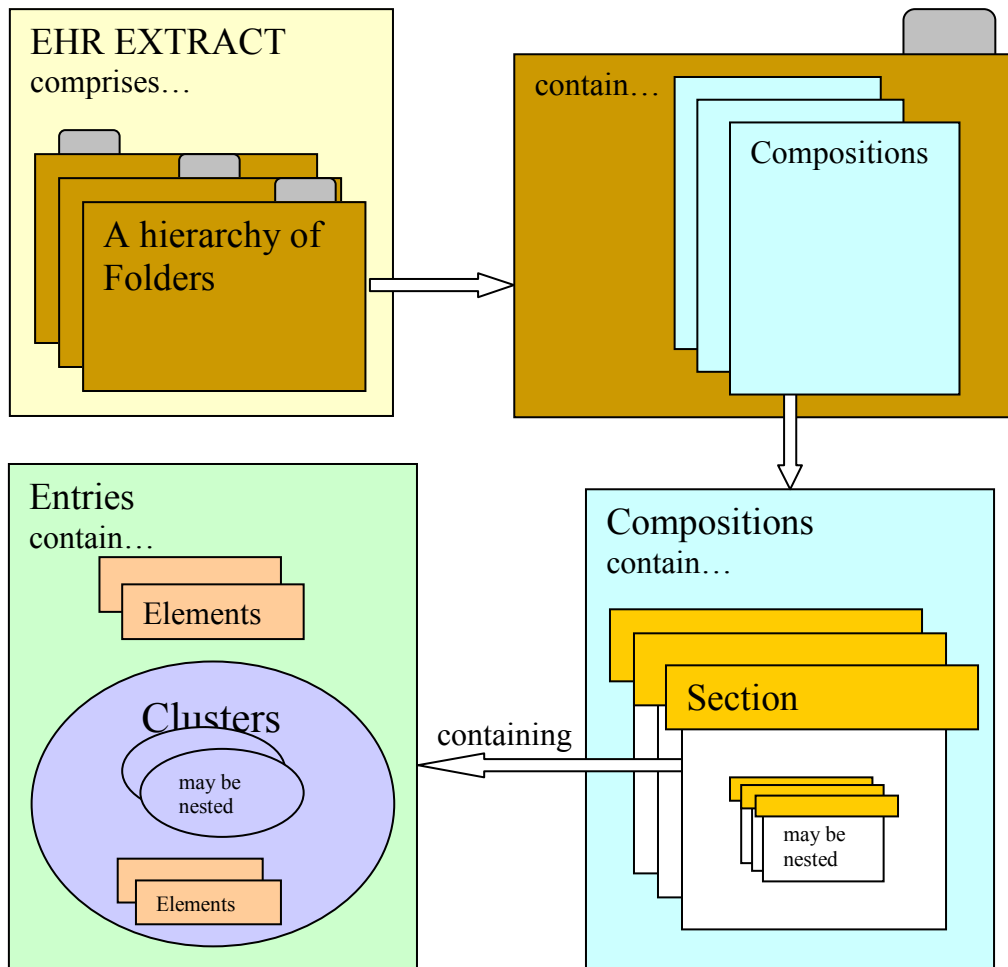


Figure 6. EHRcom hierarchy

Figure 6 gives a graphical view of the same structure, where the EHR_EXTRACT contains data of the EHR such COMPOSITION, optionally organized in a hierarchy of FOLDER. A COMPOSITION contains ENTRY, optionally contained in a hierarchy of SECTION. An ENTRY contains ELEMENTS, optionally contained in a hierarchy of CLUSTER.

4. HL7

HL7 (Health Level Seven) [HL7] is an ANSI accredited Standards Developing Organization (SDO) founded in 1987. HL7 is a not-for-profit volunteer organization that creates standards for the exchange, management, and integration of electronic healthcare information and the management, delivery, and evaluation of healthcare services.

Some HL7 aims are the development of standards that permit structured, encoded health care information to be exchanged between computer applications; develop a formal methodology to support the creation of HL7 standards from the HL7 Reference Information Model (RIM); promote the use and benefits of such standards; collaborate with other standards development organizations, standardization organizations (e.g. ANSI and ISO) and healthcare information technology users for the use of supportive

and compatible standards and to ensure that HL7 standards meet real-world requirements.

The Reference Information Model (RIM) is shown in Figure 7 [HL7 RIMa]. RIM is an essential part of the HL7 Version 3 development process. The RIM is the reference for class and attribute definitions, represents the clinical data and the connections that exist between the information carried in the fields of HL7 messages. Refined Message Information Model (RMIM) a subset of the RIM is used to express the information content for one or more related messages. The RMIM permits a more specialized information model and supports specific information constraints. RMIM is a graphical aid used to design messages and explain what each message consist of.

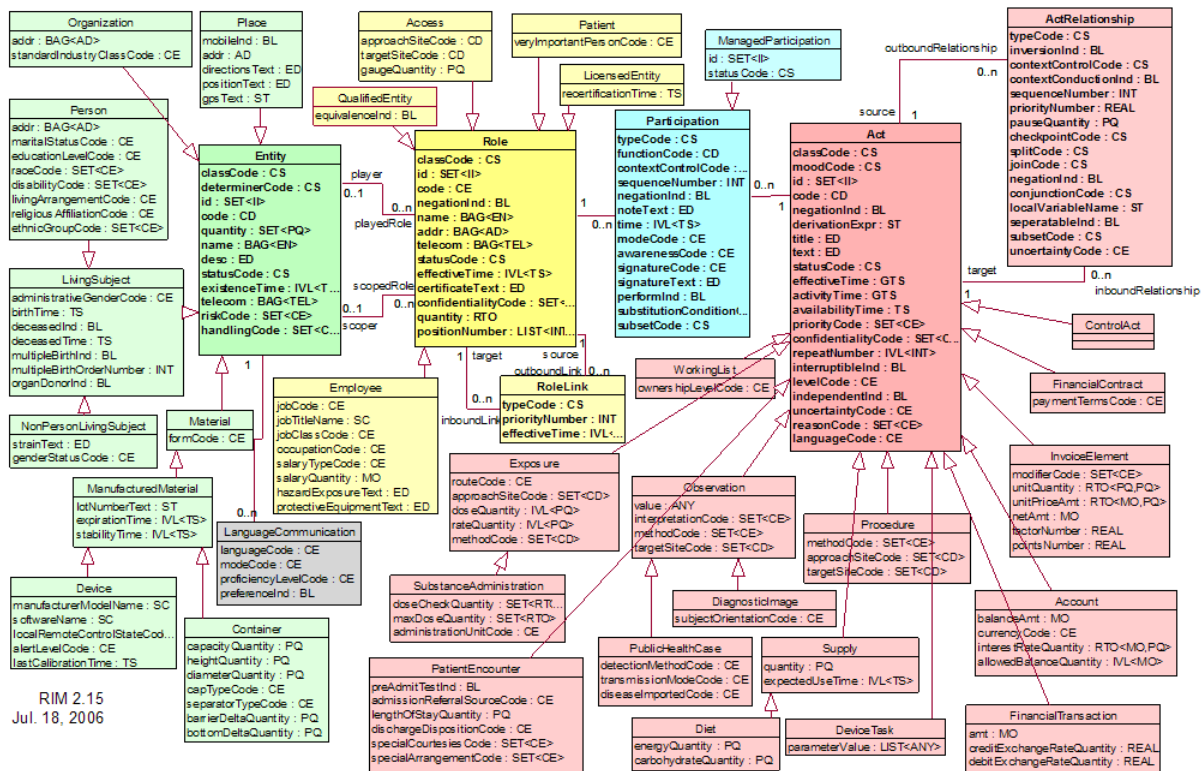


Figure 7. Reference Information Model

Inside the HL7 organization there is a Vocabulary Technical Committee, whose aim is to identify, organise and maintain vocabulary terms used in the messages. This committee is working to provide a coded vocabulary that will enable the exchange of clinical data and information with no ambiguities in the meaning of the data transferred when are used with HL7 and related standards.

4.1. HL7 Message Messaging

The HL7 standard assumes that an event in the healthcare world causes the exchange of messages between a pair of applications, this event is called *trigger event*. Therefore when an event occurs in an HL7 system, a HL7 message is constructed with the

necessary data from the underlying application systems and it is sent to the requester system. [EICHEMBERG et al, 2005].

HL7 supports two message protocols, Version 2 and Version 3. Version 2 Messaging Standard is the most widely implemented standard. Version 2 (Version 2.x) not ensures interoperability between healthcare systems because it isn't precisely definite (there are some data fields which definition is vague and there are a wide number of optional data fields). There is neither a consistent view of the data that HL7 moves nor the relationships between the data structures. Version 3 [HL7 V3] addresses these issues by using a well-defined methodology. It is based in a object-oriented data model and Reference Information Model (RIM) to create messages [HL7 RIMa, HL7 RIMb].

HL7, initially, recommends the use of XML as an alternative for HL7 V2.3.1 messages. However, in 2000, HL7 membership ratified Version 1 of the Clinical Document Architecture, which defines a XML architecture for exchange of clinical documents. The encoding is based on XML DTDs included in the specification and its semantics are defined using the HL7 RIM (Reference Information Model) and HL7 registered coded vocabularies. In January 2005, the HL7 Membership ratified Release 2 of CDA. The initial release of Version 3 will use only XML encoding. HL7 actively participates in and supports the W3C (World Wide Wb Consortium) , the organization responsible for the development of XML.

4.2. Clinical Document Architecture

The HL7 Clinical Document Architecture (CDA) is a document markup standard that specifies the structure and semantics of "clinical documents" for the purpose of exchange [HL7 CDAa] [HL7 CDAb]. CDA documents are encoded in Extensible Markup Language (XML). They derive their meaning from the HL7 Reference Information Model RIM and use the HL7 Version 3 Data Types which are part of the HL7 RIM. CDA makes documents both machine-readable and human-readable. CDA documents can be transmitted in HL7 messages designed to transfer clinical documents. A CDA document can include text or multimedia content. The coded components of a CDA have a fixed value set, called vocabulary; this vocabulary includes HL7-defined concept and HL7 recognized coding systems such as LOINC or SNOMED.

A CDA document is comprised of a header, referred to as the "CDA Header", and a body (Figure 8) . The CDA Header identifies and classifies the document and provides information on authentication, the encounter, the patient, and the provider and facilitates document management. The body contains the clinical report.

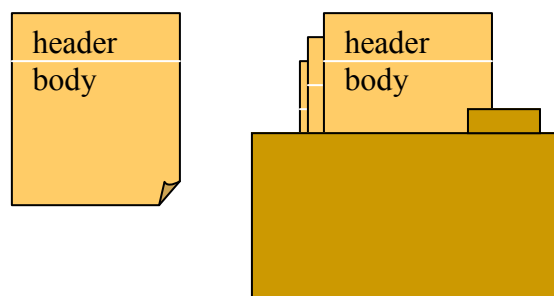


Figure 8. CDA structure

CDA is wrapped by the <ClinicalDocument> element, this contains a header between the <ClinicalDocument> and the <StructuredBody> elements. In the body each CDA section is wrapped by <section> element. Each section contains a single narrative bloc (wrapped by <text> element) which represents content to be rendered, and also it should contain CDA entries, which represent structured content provided for a computer, and finally external references (wrapped by <reference> element)

The HL7 Clinical Document Architecture is as follows:

```
<ClinicalDocument>
... CDA Header ...
<StructuredBody>
  <section>
    <text>...</text>
    <Observation>...</Observation>
    <Observation>
      <reference>
        <ExternalObservation>...</ExternalObservation>
      </reference>
    </Observation>
  </section>
  <section>
    <section>...</section>
  </section>
</StructuredBody>
</ClinicalDocument>
```

Here you have an example of a CDA component:

```
*****
Physical Exam - Lungs
*****

-->
-<component>
  -<section>
    <code code="8710-6" codeSystem="2.16.840.1.113883.6.1"
    codeSystemName="LOINC"/>
    <title>Lungs</title>
    <text>Clear with no wheeze. Good air flow.</text>
  -<entry>
    -<Observation>
      <code code="301708006" codeSystem="2.16.840.1.113883.6.96"
      codeSystemName="SNOMED CT" displayName="Respiratory sounds"/>
      <value xsi:type="CD" code="48348007"
      codeSystem="2.16.840.1.113883.6.96" codeSystemName="SNOMED CT"
      displayName="Chest clear"/>
    </Observation>
  </entry>
  -<entry>
    -<Observation negationInd="true">
      <code code="52653008" codeSystem="2.16.840.1.113883.6.96"
      codeSystemName="SNOMED CT" displayName="Respiratory sounds"/>
```

```

    <value xsi:type="CD" code="56018004" codeSystem="2.16.840.1.113883.6.96"
    codeSystemName="SNOMED CT" displayName="Wheezing"/>
  </Observation>
</entry>
</section>
</component>
-<!--

```

5. Integrating the healthcare enterprise (IHE)

Integrating the healthcare enterprise (IHE) is another organization that was created in 1998 by the Radiological the Radiological Society of North America (RSNA) [RSNA] and the Healthcare Information and Management Systems Society (HIMSS) [HIMSS]. IHE don't develop standards, otherwise it is designed to stimulate the integration and support the use existing standards, such as HL7, ASTM, DICOM, ISO, IETF, OASIS and others. In fact, it recommends and selects appropriate standards to be used in specific use cases and develops restrictions. Thus, the main aim of IHE is stimulate integration of healthcare information resources.

IHE defines a Technical Framework for the implementation of established messaging standards to achieve specific clinical goals [IHE, 2006]. This Technical Framework covers these use cases: Cardiology, Eye Care, IT Infrastructure, Laboratory, Patient Care Coordination, Patient Care Devices, Radiation Oncology, and Radiology. IHE Technical Framework have two parts, on a hand Integration Profiles, which model the business process problem and its solution, and on the other hand, Transactions, which defines the way in which current standards are used to solve the business problem defined in the Integration Profiles. Two integration profiles are, for example, Retrieve Information for Display (RID) and Cross-Enterprise Document Sharing (XDS) which address how to access to EHR in various formats. Also, IHE has developed a security framework for protecting the confidentiality, authenticity and integrity of patient care data. IHE should have an increase in medical market because it is strongly supported by the industry

5.1. Retrieve Information for Display (RID)

RID is defined as a Web Service by providing WSDL (Web Service Description Language) description with binding to HTTP GET. RID provides simple and rapid read-only access to patient-centric clinical information that is located outside the user's current application but is important for better patient care. For example information such as allergies, current medications, summary of reports, etc. for presentation to a clinician. It also supports access to existing persistent documents in different formats such as CDA (Level 1), PDF, JPEG, etc. Figure 9 shows the actors involved in this Profile and the transactions between actors.

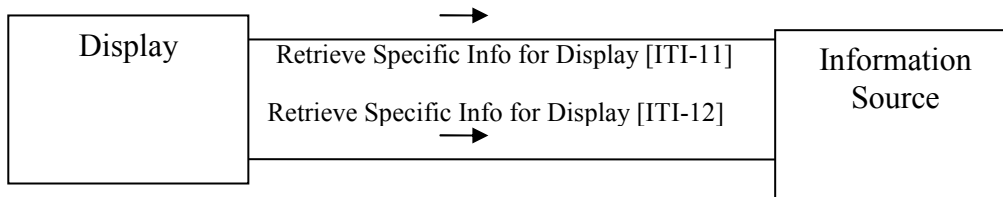


Figure 9. Retrieve Information for Display Diagram [IHE b, 2006]

5.2. Cross-Enterprise Document Sharing (XDS)

IHE defines XDS as “a set of attested clinical information (structured or not) which form an element of a patient record to be shared”. The XDS Profile facilitates the registration, distribution and access across health enterprises of patient electronic health records. XDS stores healthcare documents in an ebXML registry/repository to facilitate sharing of the documents. The document can contain any type of information in any standard format such as simple text, formatted text (HL7 CDA Release One), and images (DICOM), or structured and vocabulary coded clinical information.

6. DICOM

The American College of radiology (ACR) and the National Electrical Manufacturers Association (NEMA) [NEMA] recognized the necessity of a standard for transferring images and associated information. Both organizations decided, in 1983, to develop a standard to promote communication of digital image information. The purpose was to facilitate the development of picture archiving, and allow the creation of diagnostic information data bases that can be accessed by a wide variety of devices distributed geographically. The result was DICOM; Digital Imaging and Communications in Medicine [DICOM, 2006]. It is an standard for the communication of images, which addresses the exchange of digital information between medical imaging equipment and other systems. DICOM is typically used inside radiology, surgery, radiotherapy, and similar departments. The emphasis in these departments is high quality viewing and image processing.

Figure 10 presents the general communication model of the standard which relate both network (online) and media storage interchange (off-line) communication. The Upper Layer Service provides independence from specific physical networking communication support and protocols and the Basic DICOM File Service provides access to Storage Media independently from specific media storage formats and file structures.

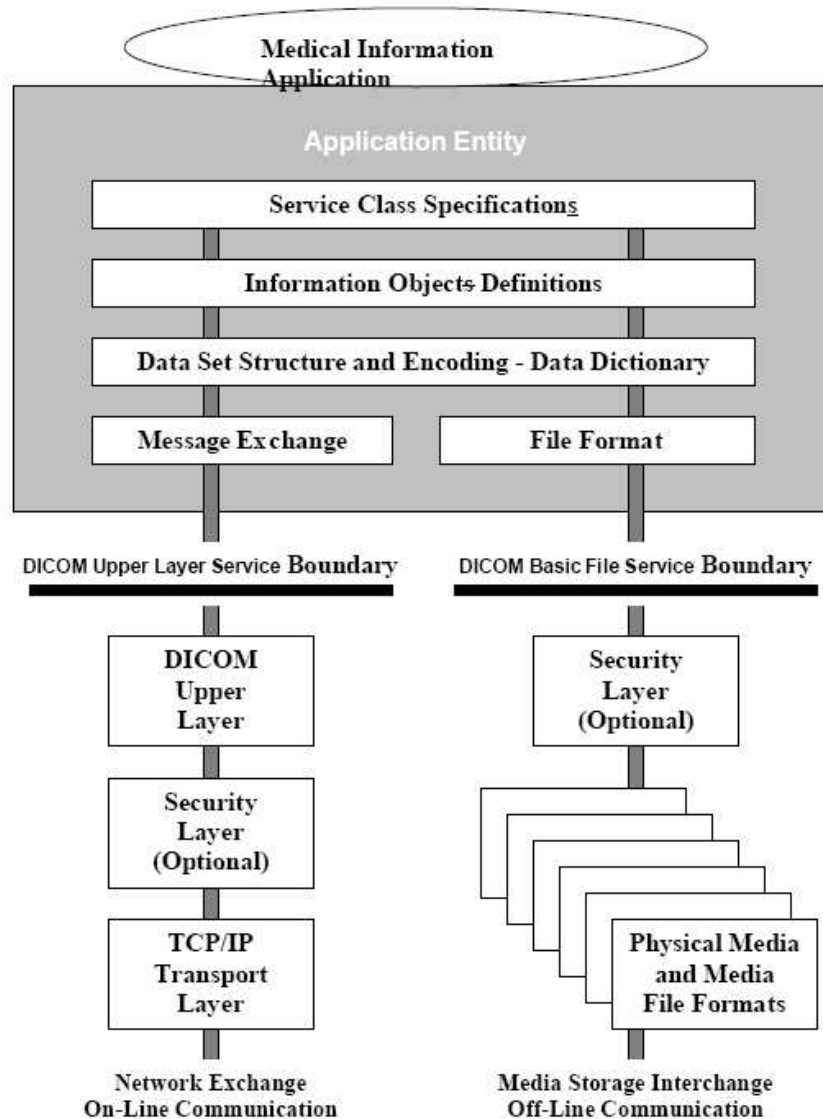


Figure 10. General Communication Model of DICOM [DICOM, 2006]

6.1. Web Access to DICOM Persistent Objects (WADO)

WADO is a part of the DICOM standard, it defines a web based service for accessing and presenting DICOM persistent objects, such as images and medical imaging reports, via HTTP or HTTPS from a Web server, using DICOM UIDs (Unique Identifiers) [WADO, 2006]. This standard is developed together by ISO TC 215 and DICOM, for this reason it is published by both organizations.

The access to the content of a data object is enabled by specifying a "link" pointing to a specific DICOM Persistent Object by means of its URL/URI and using DICOM UIDs. There are some optional fields that a web client can specify to the server, such as anonymize the DICOM object (remove patient identifying information), or request to the server that a DICOM object are represented in another format like JPEG.

6.2. DICOM Structured Reporting (DICOM SR)

DICOM Structured Reporting is an extension of the DICOM standard created in 2000 that covers medical reports and other clinical data [DICOM, 2006]. It is a general model to encode medical reports in a structured manner in a tag-based format. A structured report has a header information that is also used for DICOM images, and the content is represented by a document tree. Parent and child content items are related to each other by a set of relationships. The content items of the tree contain some pieces of information such as a text paragraph or a reference to an image, they also include machine readable codes. These codes are taken from standard vocabularies such as SNOMED [SNOMED] or LOINC [LOINC].

7. Comparative analysis

In this survey we have described in more detail the three standards that are taking the leadership: open EHR, EHRcom and HL7 standards. Nevertheless, to have a wider perspective of Electronic Healthcare Records standards other approaches have been introduced, in particular: DICOM (WADO, DICOM SR) and IHE. Just to mention that OpenEHR has been explained because it provides some basic concepts to describe the other standards, but in this section openEHR is subsumed in EHRcom.

Initially, it is necessary to distinguish the scope of the standards. Some EHR standards only specify content structure and others only specify access services. EHRcom and DICOM SR specify content format and communication protocol, while HL7 CDA only specify content format, and ISO WADO, IHE RID and IHE XDS only specifies communication protocol [EICHEBERG et al, 2005].

The content format of EHRcom, DICOM SR and HL7 CDA standards is very similar [EICHEBERG et al, 2005]. All of these standards can be used to store persistent structured documents. For these standards, the persistent document is the basic unit in which information is assembled, stored and communicated. EHRcom differs from other standards because an EHRcom extract contains one or more Compositions (equivalent to documents) in nested folders.

EHRcom, DICOM SR and HL7 CDA standards also can contain multimedia data (images, movies, signals) and references to this multimedia data in the health care record. All of the content standards support the two level modelling concept, they use a simple reference model and a set of constraint rules that describes how clinical observations can be expressed without ambiguities. EHRcom's constraint rules are archetypes and HL7 CDA and DICOM SR are designed to have templates. At the moment, HL7 has a special interest group for templates but HL7 CDA templates are undefined, although the aim is that templates have to be compatible with archetypes. In fact, ADL archetypes can be written against any UML model, and it would be possible to write archetypes directly against the HL7v3 RIM (Reference Information Model), and also the CDA specification. All of the standards have the aim of specifying a library of archetypes or templates. Except DICOM the other standards don't have a largely defined library.

To end with content-based issues, CDA gives some recommendations on how documents are to be structured are encoded to be presented to a human reader. CDA require that human readable information must be in the narrative bloc.

Concerning the communication aspect, EHRcom, ISO WADO, DICOM SR, IHE RID, and IHE XDS specify access services (communication protocols). EHRcom refers to its communication protocol ENV-13606-5, and DICOM-SR refers to DICOM Storage and Query/Retrieve Services that are used to transport DICOM SR documents and images. The basic access services in the EHR standards are query, retrieval and submission. All standards accept these three services except WADO. WADO does not allow querying for document content. It needs to be combined with other services such XDS for queries. WADO and, also, IHE RID don't support the submission of new documents to a repository. Finally, we have that in EHRcom they use the concept of EHRcom extract, which may contain one or more documents, instead of persistent document as the basic unit of information for query and retrieval but not necessarily for submission.

Some of these standards can provide security mechanisms for encryption, user credentials and access control. EHRcom not have a defined protocol, for this reason, its security features are still unknown. ISO WADO, DICOM SR, IHE RID, and IHE XDS support the encryption of the record content during the transmission, this is implemented using the Transport Layer Security (TLS) protocol. Optionally, these standards allow to transmit user credentials, this permits to the system to know who is the user that is requesting access. It can be necessary that somebody assumes the responsibility for a clinical document, for example with a signature. CDA and EHRcom don't specify explicitly the abilities to attach a digital signature to EHR documents, because this standards can be represented in XML so digital signatures are handled through the XML Signature Standard. Nevertheless DICOM-SR standard explicitly specifies who to attach a digital signature.

In conclusion, it can be said that content format standards are quite similar. There are some differences, mainly, in the progress of standardization. While DICOM SR already has a library of templates, EHRcom and CDA are in the beginning of the process. The EHR access service standard differs of each others for scope and progress achieved in standardization process. DICOM is the most stable standard but it is focused to medical imaging and that, probably, make it unsuccessful in the general EHR market. The access services for EHRcom and CDA are still undefined, although they are under development.

References

[Beale, 2002]

Thomas Beale. Archetypes: Constraint-based domain models for future-proof information systems. In: Eleventh OOPSLA Workshop on Behavioral Semantics: Serving the Customer (Seattle, Washington, USA, November 4, 2002). Edited by Kenneth Baclawski and Haim Kilov. Northeastern University, Boston, 2002, pp. 16-32.

[Beale et al, 2005]

Beale T., Goodchild A., Tun Z., Austin T., Kalra D., Lea N. Lloyd D., The Archetype Object Model. Rev. 2.0. The OpenEHR Foundation. 2005.

[Beale et al, 2006]

Beale T., Heard S., Kalra D., Lloyd D. Reference Model, The openEHR EHR Information Model. Rev. 5.0. The OpenEHR Foundation. 2006.

[Beale and Heard, 2005]

Beale T., Heard S. Archetype definitions and principles. Rev. 0.6. The OpenEHR Foundation. 2005.

[Beale and Heard, 2006]

Beale T., Heard S. The archetype definition language (ADL). Rev. 1.3.1. The OpenEHR Foundation. 2006.

[CEN/TC 251]

European Committee for Standardization-Technical Committee on Health Informatics. <http://www.centc251.org/>.

[CEN prEN 13606-1, 2006]

CEN prEN 13606-1 Health Informatics-Electronic health record communication-Part 1: Reference Model. Draft European Standard for CEN Enquiry prEN 13606-1, European Committee for Standardization. Brussels, Belgium. 2006.

[Cohen, 2001]

Cohen S., Shabo A., Electronic Health Record (EHR) Standards Survey, 2001 <http://www.haifa.il.ibm.com/projects/software/imr/papers/EHRSurvey.pdf> , 2006

[DICOM, 2006]

Digital Imaging and Communications in Medicine (DICOM) Part 1: Introduction and Overview, PS 3.1-2006, National Electrical Manufacturers Association, 2006

[EICHEMBERG et al, 2005]

Marco Eichelberg, Thomas Aden, Jörg Riesmeier, Asyman Dogac and Gorkce B. Laleci, A survey and analysis of Electronic Healthcare record standards, ACM Computer Surveys, Vol 37, No. 4, pp. 277-315, 2005

[IHE, 2006]

IHE. IT infrastructure technical framework Rev. 2.0. Tech. rep., Integrating the Healthcare Enterprise. http://www.ihe.net/Technical_Framework/index.cfm., 2006

[IHE b, 2006]

IHE Cardiology Technical Framework, Volume I, Integration Profiles, Rev. 2.1, June 2006

[HIMSS]

HIMSS. Healthcare Information and Management Systems Society. <http://www.himss.org/>.

[HL7]

Health Level 7 (HL7) www.hl7.org

[HL7 CDAA]

HL7. Clinical Document Architecture, Release 2.0, 2004

[HL7 CDAB]

HL7. Clinical Document Architecture Framework Version 1.0. August, 2000

[CDA]

CDA sample XML instance <http://xml.coverpages.org/CDA-ReleaseTwoSample200403.xml>

[HL7 RIMa]

HL7. HL7 Reference Information Model, Version 2.15, http://www.hl7.org/library/data-model/RIM/modelpage_mem.htm

[HL7 RIMb]

HL7. HL7 Reference Information Model, <http://www.hl7.org/library>, 2001

[HL7 V3]

George W Beeler, Stan Huff, Wesley Rishel, Abdul-Malik Shakir, Mead Walker, Charlie Mead, Gunther Schadow. Message Development Framework. V3.3, <http://www.hl7.org/library/mdf99/mdf99.pdf>, 1999

[Iakovidis1998]

Iakovidis, I. 1998. Towards personal health records: Current situation, obstacles and trends in implementation of Electronic Healthcare Records in Europe. *Int. J. Medical Informatics* 52(1), 105-115

[ISO/DTR 20514]

ISO/DTR 20514. Health informatics - electronic health record: definition, scope and context. Schloeffel P, ed. Fourth Draft, March 2004.

[LOINC]

LOINC, Logical Observation Identifiers Names and Codes. <http://www.loinc.org/>.

[NEMA]

National Electrical Manufacturers Association, <http://medical.nema.org/>

[openEHR]

openEHR Community <http://www.openehr.org/>

[RSNA]

RSNA. Radiological Society of North America. <http://www.rsna.org/>.

[SNOMED]

SNOMED (The Systematized Nomenclature of Medicine) Clinical Terms.

http://www.snomed.org/snomed_txt.html

[WADO, 2006]

Digital Imaging and Communications in Medicine (DICOM), Part 18: Web Access to DICOM Persistent Objects (WADO), 2006

[WHO, 2006]

World Health Organization. <http://www.who.int/classifications/icd/en/>