A Survey and Analysis of Electronic Healthcare Record Standards

MARCO EICHELBERG and THOMAS ADEN
Oldenburger Forschungs- und Entwicklungsinstitut für Informatik-Werkzeuge und -Systeme (OFFIS)

and

ASUMAN DOGAC and GOKCE B. LALECI
Middle East Technical University (METU)

Medical information systems today store clinical information about patients in all kinds of proprietary formats. To address the resulting interoperability problems, several Electronic Healthcare Record standards that allow to structure the clinical content for the purpose of exchange are currently under development. In this article, we present a survey of the most relevant Electronic Healthcare Record standards, examine the level of interoperability they provide and assess their functionality in terms of content structure, access services, multimedia support and security. We further investigate the complementarity of the standards and assess their market relevance.

Categories and Subject Descriptors: J.3 [**Computer Applications**]: Life and Medical Sciences—*Medical Information Systems*; H.2.4 [**Database Management**]: Systems—*Distributed databases*; H.3.4 [**Information Storage and Retrieval**]: Systems and Software—*Distributed systems*

General Terms: Standardization, Design, Performance

Additional Key Words and Phrases: eHealth, Electronic Healthcare Record Standards, Interoperability

1. INTRODUCTION

The Electronic Health Record (EHR) is defined as digitally stored healthcare information about an individual's lifetime with the purpose of supporting continuity of

Authors' addresses: M. Eichelberg and T. Aden, Kuratorium OFFIS e.V., Escherweg 2, 26121, Oldenburg, Germany, e-mail: {Marco.Eichelberg, Thomas.Aden}@OFFIS.de; A. Dogac and G. B. Laleci, Dept. of Computer Eng., Software R&D Center, Middle East Technical University (METU), 06531, Ankara, Turkey, e-mail: {asuman, gokce}@srdc.metu.edu.tr.

This work is supported by the European Commission through the RIDE project and in part by the Scientific and Technical Research Council of Turkey (TÜBİTAK), Project No: EEEAG 104E013

care, education and research. It includes information such as observations, laboratory tests, diagnostic imaging reports, treatments, therapies, drugs administered, patient identifying information, legal permissions, and allergies. Currently, this information is stored in all kinds of proprietary formats through a multitude of medical information systems available on the market. Typical formats include relational database tables; structured document-based storage in various formats and unstructured document storage such as digitised hardcopies maintained in a classical document management system. This results in a severe interoperability problem in the healthcare informatics domain.

Making EHRs interoperable will contribute to more effective and efficient patient care by facilitating the retrieval and processing of clinical information about a patient from different sites. Transferring patient information automatically between care sites will speed delivery and reduce duplicate testing and prescribing. Automatic reminders will reduce errors, improve productivity, and benefit patient care. Furthermore, one of the prominent research directions in the medical field is about using genomics data for improving health knowledge and processes for prevention, diagnosis, treatment of diseases and personalisation of health care. This also necessitates the interoperability of biomedical information and the EHRs.

Interoperability has been a challence in the IT domain for a long time now and several approaches have been developed at the syntactic level such as ODBC data gateways, message queues and interface engines, software adapters and more recently, Web services. While such developments have been very useful, providing interoperability at the schema and data level is still a very difficult problem. The problem gets more complicated in the healthcare domain since clinical information itself is very complex: there are more than 300.000 clinical concepts and many coding systems.

To address the EHR interoperability problem, there are several standards currently under development such as the Health Level 7 (HL7) Clinical Document Architecture (CDA) [HL7CDA Release 2.0 2004], CEN 13606 EHRcom [EHRcom 2004] and openEHR [OpenEHR]. These standards aim to structure and markup the clinical content for the purpose of exchange. There is also an industry initiative called "Integrating the Healthcare Enterprise" (IHE) [IHE IT-IP] which specified the "Cross Enterprise Document Sharing (XDS)" Integration Profile [IHE-XDS 2004] for this purpose. The basic idea of IHE XDS is to store healthcare documents in an ebXML registry/repository architecture to facilitate their sharing.

Given this important interoperability problem in the healthcare domain and given many standards and industry initiatives addressing the problem, this paper intends to analyse the prominent EHR standards to shed some light on the following questions:

—*The level of interoperability support*: Does the EHR provide structured content suitable for automated processing? Does it specify content distribution rules?

—*Functionality*: Does the standard allow for an explicit retrieval of records (or parts thereof) for a specific patient, based on an incoming request? Can it contain multimedia data? What kind of security mechanisms are supported for accessing healthcare records?

—*Complementarity*: Since not all the standards provide all the necessary features, is

it possible to combine them in a complementary way? Do the standard initiatives affect one another?

—*Market relevance*: Is the standard accepted in the marketplace? Are there commercial implementations available or any signs of uptake by industry?

The paper is organized as follows: Section 2 summarizes the GEHR/openEHR initiative. This EHR specification introduces an important concept, called the "archetype" which is a reusable, formal expression of a distinct, domain-level concept such as "blood pressure", "physical examination", or "laboratory results", expressed in the form of constraints on data whose instances conform to some reference model [Beale and Heard]. This concept has been adopted or developed independently by other prominent EHR standards. Section 3 discusses the CEN EN 13606 EHRcom standard by the European Standardization organization of Health Informatics. Section 4 introduces the HL7 standard and describes the HL7 document markup standard called the Clinical Document Architecture (CDA). Although CDA is not an EHR standard as such, it forms an important component of an EHR. In Section 5, the de-facto standard for medical image communication, namely, Digital Image and Communications in Medicine (DICOM) is presented. An important standard providing Web access to DICOM Persistent Objects, called WADO, is also given in this section. The EHR related standards produced by "ISO/TC 215 Health Informatics" are covered in Section 6. Section 7 describes the "Integrating the Healthcare Enterprise" initiative. Section 8 briefly introduces the Medical Markup Language (MML). Section 9 contains a comparison and evaluation of the presented standards. Finally, Section 10 concludes the paper.

2. THE GEHR/OPENEHR INITIATIVE

The GEHR/openEHR initiative was started in 1992 as an EU research project, called "Good European Health Record", in the third framework programme. The initiative was later continued under the name "Good Electronic Health Record" with strong participation from Australia. Currently it is maintained by the openEHR Foundation, a non-profit organisation defining itself as "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 most noteworthy concept introduced by GEHR/openEHR is the "archetype" concept [Beale 2002]. This approach uses a two-level methodology to model the EHR structure. In the first level, a generic reference model that is specific to the healthcare domain but still very general is developed. This model typically contains only few classes (e. g. role, act, entity, participation) and must be stable over time. In the second level, healthcare and application specific concepts such as blood pressure, lab results etc. are modelled as archetypes, that is, constraint rules that specialise the generic data structures that can be implemented using the reference model. As an example, a constraint may restrict a generic "Observation" class to, say, "Blood Pressure" archetype.

An archetype definition basically consists of three parts: descriptive data, constraint rules and ontological definitions. The descriptive data contains a unique identifier for the archetype, a machine-readable code describing the clinical concept

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

Fig. 1. The ADL definition of "Complete Blood Count" Archetype [CBC Archetype]

modeled by the archetype and various metadata such as author, version, and purpose. It also states whether an archetype is a specialisation of another archetype. The constraint rules are the core of the archetype and define restrictions on the valid structure, cardinality and content of EHR record component instances complying to the archetype. The ontological part defines the controlled vocabulary (that is, machine readable codes) that can be used in specific places in instances of the archetype. It may contain language translations of code meanings and bindings from the local code values used within the archetype to external vocabularies such as SNOMED [SNOMED] or LOINC [LOINC]. It may also define additional constraints on the relationship between coded entries in the archetype based on the code value.

A formal language for expressing archetypes has been introduced by the openEHR initiative which is called Archetype Definition Language (ADL) [ADL 1.2]. Figure 1 presentes an example archetype definition expressed in ADL describing a part of the "Complete Blood Count" concept. The complete ADL definition can be found in [CBC Archetype]. Here the "OBSERVATION" class from the reference information model is restricted to create "Complete Blood Count" archetype, by constraining its CODED_TEXT value to "ac0001" term, (the term ac0001 is defined to be "complete blood count" in the constraint_definitions part of the ADL, and declared to be equivalent to the LOINC code 700-0 in the term bindings part), and by defining its content to be a list of "Haemoglobin", "Haematocrit" and "Platelet

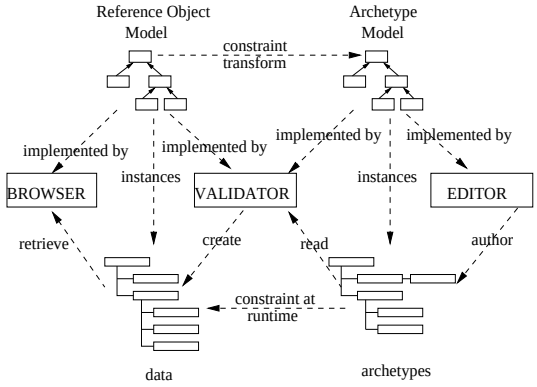

Fig. 2. openEHR Archetype Methodology [Beale 2002]

Count" test result elements.

As already mentioned, the archetype model or constraint language is closely coupled to the reference model since archetypes define constraints on the reference model. Similarly, each data instance in the EHR is closely coupled to one archetype which defines the constraints the data has to comply to in addition to the rules of the reference model. In this approach, an EHR system needs to offer three building blocks that are specific to the archetype approach: an editor for creating and maintaining archetypes, a validator that enforces the constraints at runtime and a browser component that allows for an optimised display of specific archetypes as shown in Figure 2.

Data entry into an EHR system is checked at runtime against the constraints defined in the archetype for the concept to be stored, however, the information system itself just maps the entered data onto the reference model, which means that the data structures (e. g. relational tables) of the information system are quite stable even if clinical concepts change over time. In Figure 3, we depict an example in order to provide better insight into how the "archetype" concept helps an EHR system to accommodate changes without modifying the information system data structures. The two archetype instances in this figure, namely, the "Blood Pressure Archetype" instance and the "Demographics Archetype" instance express the clinical domain concepts independent of how the data is actually stored in a relational database. In this way, it is possible to modify the healthcare concepts, that is the archetypes and their instances without changing the underlying relational tables.

The term "archetype" also hints to the concept of a library of archetypes that can be used as a starting point for developing a particular EHR system, as opposed to a "development from scratch" approach. The openEHR framework includes a reference information model, a formal language for expressing archetypes, called Archetype Definition Language (ADL) [ADL 1.2], an archetype library, implementation technology specifications (XML schemas, IDL specifications etc.) and a collection of open source implementations of the openEHR specifications.

It should be noted that HL7 CDA Templates and DICOM Structured Reporting

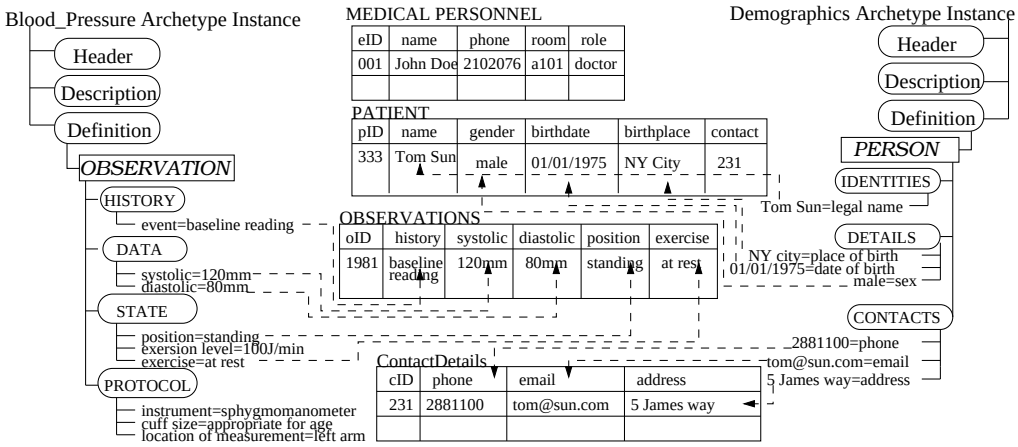

Fig. 3. An Example Association between Archetypes and the Information System Data Structures

Templates are conceptually very similar to archetypes.

3. CEN/TC 251 AND ENV/EN 13606 EHRCOM

CEN/TC 251 [CEN TC251] is the technical committee on Health Informatics of the European Committee for Standardization. Its mission is to achieve compatibility and interoperability between independent health systems and to enable modularity by means of standardisation. This includes requirements on health information structure to support clinical and administrative procedures, technical methods to support interoperable systems as well as requirements regarding safety, security and quality.

The CEN Pre-standard ENV 13606:2000 "Electronic Healthcare Record Communication" [CEN ENV 13606 2000] is a message-based standard for the exchange of electronic healthcare records. The standard defines an EHR information model, called the "extended architecture" since it is an extension of the earlier pre-standard ENV 12265 [CEN ENV 12265]. It also defines a list of machine-readable domain terms that can be used to structure EHR content, a method of specifying "distribution rules", that is, rules under which certain EHR content may be shared with other systems and, finally, request and response messages that allow systems to exchange subsets of an EHR. ENV 13606 does not attempt to specify a complete EHR system, instead it focuses on the interfaces relevant for a communication between EHR systems.

ENV 13606 was arguably the first fully implementable EHR standard, and subsets of it were implemented in a number of EHR projects in the UK, Denmark, the Netherlands, Sweden, and Norway. However, none of these projects used the complete ENV 13606 specification and implementation experience showed a number of weaknesses in the standard that limited its usefulness and market uptake. The single-level modelling approach made the information model extremely complex, with lots of optionality and a level of abstraction that made the model quite

difficult to comprehend and implement.

In 2001, CEN/TC 251 decided to revise ENV 13606 into a full European Standard, taking into account the existing implementation experience and to adopt the openEHR archetype methodology. EN 13606, called EHRcom, will be a five-part standard consisting of:

(1) The Reference Model,
(2) Archetype Interchange Specification,
(3) Reference Archetypes and Term Lists,
(4) Security Features, and
(5) Exchange Models.

CEN/TC 251 also plans to introduce EHRcom into ISO/TC 215 as the basis for an international EHR standard. Currently, however, only the reference model (EN 13606-1) is stable, whereas parts 2 through 5 are still working drafts. Therefore, the following discussion mostly focuses on the reference model as defined in [EHRcom 2004].

The EHRcom reference model consists of four packages: *Extract*, *Demographics*, *Access Control* and *Message* which together describe the aspects of an EHR that are relevant for communication of EHR extracts between information systems:

— *The Extract package* defines the root class of the reference model ("EHR_EXTRACT") and the data structures for EHR content.
— *The Demographics package* provides a minimal data set to define the various persons, software agents, devices and organisations that are referenced within the EHR extract. These data structures are based on the General Purpose Information Components (GPICs) defined in [CEN prEN 14822-1 2003].
— *The Access Control package,* which is under development as EN 13606-4, will define a representation for EHR access policies (such as consents for disclosure).
— *The Message package,* which is under development as EN 13606-5, will define the attributes required to communicate the EHR extract to a requesting process via a message or other serialised form. This part of the EHRcom standard will also include an HL7 Domain Message Information Model (D-MIM) corresponding to the EHRcom reference model, i. e. allowing to define HL7 version 3 messages to be used for communication of EHR extracts.

Figure 4 shows the logical building blocks of EHR content. The top level is a directory of possibly nested folders for a patient, allowing for a high-level organisation of the EHR, for example, per episode or per clinical speciality. Folders contain zero or more "compositions" by reference. A composition (which roughly corresponds to one clinical document) may contain sections with section headers and entries which consist of elements or clusters of elements. Each element has a single value of a single data type. Content in the EHR extract is always added or replaced as a complete composition – versioning, ownership and audit trail in EHRcom are based on the composition.

The second important building block for EHRcom is the archetype methodology. As described in Section 2, archetypes allow to describe specific clinical concepts such as blood pressure or ECG measurements, as constraint rules that restrict the

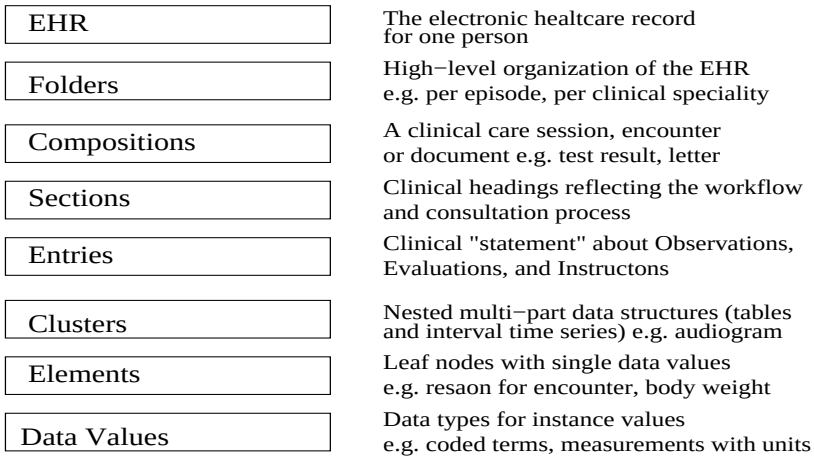

Fig. 4. Logical Building Blocks of EHRcom [Kalra 2004]

possible types, relationships and values of the record components in an EHRcom composition, or a part thereof. EN 13606-2 will define an archetype description language (ADL), that is, a formal language that is related to the EHRcom reference model. Archetypes expressed in this language will be convertible to HL7 Refined Message Information Models (R-MIMs) and Common Message Element Types (CMETs). It is intended to harmonise the EHRcom archetype concept with the HL7 Clinical Document Architecture (CDA) and HL7 Templates.

EN 13606-3 will contain a library of archetypes for various purposes, similar in principle to the library of Structured Reporting templates in part 16 of the DICOM standard. This work in EHRcom will most likely be based on the CEN General Purpose Information Components (GPICs) defined in [CEN prEN 14822-1 2003].

At this time it is not yet possible to make any statement on implementations or market acceptance of EHRcom since only the reference model is stable and EHRcom parts 2-5 are still under development. However, the harmonisation with HL7 and openEHR as well as the improvement of the architecture based on the lessons learned with ENV 13606 will certainly improve usability and acceptance over the 1999 pre-standard.

4. HL 7

Founded in 1987, HL7 (Health Level Seven) [HL7] is a not-for-profit, ANSI-accredited Standards Developing Organization that provides standards for the exchange, management and integration of data that supports clinical patient care and the management, delivery and evaluation of healthcare services.

The standard is developed with the assumption that an event in the healthcare world, called the *trigger event*, causes the exchange of messages between a pair of applications. When an event occurs in an HL7 compliant system, an HL7 message is prepared by collecting the necessary data from the underlying application systems and it is passed to the requestor, usually as an EDI message. For example, a trigger event, called ADT, can occur when a patient is admitted to a hospital and

this may cause the data about the patient to be collected and sent to a number of other application systems.

Currently, there are two message protocols supported by HL7: Version 2 and Version 3 which was approved in 2004 by the American National Standards Institute [ANSI]. HL7 Version 2 Messaging Standard is the most widely implemented standard for healthcare information in the world today. However, being HL7 Version 2 compliant does not imply direct interoperability between healthcare systems. This stems from the fact that Version 2 messages contain many optional data fields. This optionality provides great flexibility, but necessitates detailed bilateral agreements among the healthcare systems to achieve interoperability. To remedy this problem, HL7 is developing Version 3 [HL7V3] which is based on an object-oriented data model, called Reference Information Model (RIM) [HL7RIM].

Up to the current Version 2.5 [HL7v2.5 2000], the scope of the HL7 standard was limited to the exchange of messages between medical information systems. Starting with Version 3.0, a document markup standard, called the Clinical Document Architecture (CDA) is proposed as presented in Section 4.1.

4.1 HL7 Clinical Document Architecture (CDA)

CDA, previously called Patient Record Architecture (PRA), defines structure and semantics of medical documents for the purpose of exchange. CDA documents are encoded in Extensible Markup Language (XML) and 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.

Table I. Evolution of CDA "levels" from Release One to Release Two

CDA Release One	*CDA Release Two*
CDA Level One	The unconstrained CDA specification
CDA Level Two	The CDA specification with section-level templates applied
CDA Level Three	The CDA specification with entry-level templates applied

CDA is organized into three levels as shown in Table I where each level iteratively adds more markup to clinical documents, although the clinical content remains constant at all levels. "Level One" [HL7CDA Release 1.0 2000], which is already a standard, focuses on the content of narrative documents with high-level context such as parties, roles, dates and time, places and structural organisation of headings. It consists of two parts, the CDA Header and the CDA Body, which are based on the HL7 data types. The document header is derived from RIM and unambiguously defines the semantics of each entry in the document. The body contains the clinical report, and can be either an unstructured text, or can be comprised of nested containers such as sections, paragraphs, lists, and tables through structured markup. Hence there is no semantics in Level One body; it offers interoperability only for human-readable content. In fact, CDA Level One describes a kind of HTML document with a standardised header that contains additional information on the document.

```
<section>
  <code code="8709-8" codeSystem="2.16.840.1.113883.6.1" codeSystemName="LOINC"/>
  <title>Skin Exam</title>
    <text>Erythematous rash, palmar surface, left index finger.
      <renderMultiMedia referencedObject="MM2"/> </text>
  <entry>
    <Observation>
      <code code="106076001" codeSystem="2.16.840.1.113883.6.96"
        codeSystemName="SNOMED CT" displayName="Skin finding"/>
      <value xsi:type="CD" code="271807003"
        codeSystem="2.16.840.1.113883.6.96"
        codeSystemName="SNOMED CT" displayName="Rash"/>
      <targetSiteCode code="48856004"
        codeSystem="2.16.840.1.113883.6.96" codeSystemName="SNOMED CT"
        displayName="Skin of palmer surface of index finger">
        <qualifier>
          <name code="78615007" codeSystem="2.16.840.1.113883.6.96"
            codeSystemName="SNOMED CT" displayName="with laterality"/>
          <value code="7771000" codeSystem="2.16.840.1.113883.6.96"
          codeSystemName="SNOMED CT" displayName="left"/>
        </qualifier> </targetSiteCode>
      <entryRelationship typeCode="SPRT">
        <RegionOfInterest MMID="MM2">
          <id root="10.23.4567.4489"/>
          <code code="ELLIPSE"/> <value>3 1 3 7 2 4 4 4</value>
          <entryRelationship typeCode="SUBJ">
            <ObservationMedia> <id root="10.23.4567.345"/>
              <value xsi:type="ED" mediaType="image/jpeg">
              <reference value="lefthand.jpeg"/> </value>
            </ObservationMedia> </entryRelationship>
        </RegionOfInterest> </entryRelationship> </Observation>
  </entry>
</section>
```

Fig. 5. Part of a CDA Document Body [HL7CDA Release 2.0 2004]

Level Two CDA, which is a draft standard, models the fine-grained observations and instructions within each heading through a set of RIM Act classes. With Level Two, it will be possible to constrain both structure and content of a document by means of a template and thereby increase interoperability since the receiver "knows what to expect". However, a completely structured document where the semantics of each information entity is specified by a unique code will only be possible with "Level Three" providing for arbitrary machine processing.

Although the current draft of the CDA Release Two standard [HL7CDA Release 2.0 2004] does not distinguish these three levels any more (Table I), the basic architecture with structured documents of different granularity remains. This approach is intended to facilitate the migration from current free text reports to more structured CDA documents.

In order to provide further insight to CDA Level Three, a part of the body of a CDA document is provided in Figure 5. The narrative block in the <text> element is solely for human consumption whereas the rest of the document is both for human consumption and for automated processing since it is encoded in XML

and contains coded elements. For example, in Figure 5, the "Section" element has been coded with Logical Observation Identifiers Names and Codes (LOINC) [LOINC] and "Observation" element has been coded with SNOMED Clinical Terms [SNOMED] to be "Skin finding" and through these codes and markup, it is machine processable. Furthermore, document-level, section-level and entry-level templates can be used to constrain the generic CDA specification.

Unlike other standards HL7 CDA does not specify services or protocols that are used to exchange a document. From the perspective of HL7 messages, a CDA document is just a multimedia object than can be exchanged as a MIME (Multipurpose Internet Mail Extensions) package.

Many national and international pilot projects use HL7 CDA Release One as a format for clinical documents [HL7 CDA Implemantations]. Also commercial products implementing CDA are starting to become available. Since medical reports are currently mainly stored in clinical information systems which already use the HL7 standard, vendors already have experience with HL7 and are likely to be aware of the opportunities offered by CDA.

Strictly speaking, the HL7 Clinical Document Architecture (CDA) is not an EHR standard since it only defines parts of an EHR architecture. However, the CDA forms an important component of an EHR and is currently being harmonised with the equivalent structures in EN 13606 and openEHR. CDA is a subset of the EN 13606 Reference Model and EN 13606 will be compliant with CDA Release Two.

5. DIGITAL IMAGE AND COMMUNICATIONS IN MEDICINE (DICOM)

DICOM (Digital Image and Communications in Medicine) [DICOM] is known as the de-facto standard for medical image communication. The standard defines data structures and services for the vendor independent exchange of medical images and related information. It is being developed by medical industry and medical professional organisations under the umbrella of the National Electrical Manufacturers Association (NEMA) [NEMA].

In this section two DICOM based EHR standards, namely, Web Access to DICOM Persistent Objects (WADO) and DICOM Structured Reporting are covered.

5.1 Web Access to DICOM Persistent Objects (WADO)

Web Access to DICOM Persistent Objects (WADO) is a joint effort of DICOM [DICOM Supplement 85 2004] and ISO [ISO/DIS 17432 2004] and, therefore, it is published by both organisations. The standard defines a Web-based service that can be used to retrieve DICOM objects (images, waveforms and reports) via HTTP or HTTPS from a Web server. A query mechanism is not supported. The Web client has to specify the DICOM object to be retrieved by its unique identifiers for the study, series and instance level of the hierarchical DICOM information model. In addition to these required fields, there are a number of options that can be passed to the server. For example, the client can ask the server to anonymise the DICOM object before transmission, that is, to remove any patient identifying information from the data. The client can also request the server to convert the DICOM object to a different, presentation-ready format, for example, JPEG for images and HTML for reports. The standard specifies which formats (MIME types) are required and which are optional for a server. For non-DICOM images the client

has also the choice to select the size of the resulting image or a region of interest, the frame number in case of multi-frame images, the image quality in case of lossy compression or the parameters of the value of interest transformation (window centre and width). Figure 6 shows an example URL (Uniform Resource Locator) for retrieving a DICOM Structured Reporting document in HTML format.

```
https://192.168.1.2/RetrieveDocument?requestType=WADO
&studyUID=1.2.250.1.59.40211.12345678.678910
&seriesUID=1.2.250.1.59.40211.789001276.14556172.67789
&objectUID=1.2.250.1.59.40211.2678810.87991027.899772.2
&contentType=text%2Fhtml
&charset=UTF-8
```

Fig. 6. Example WADO URL for retrieveing a DICOM Structured Report

Although a WADO server is required to return any DICOM Structured Reporting document in HTML format, if the client does not make any choice on the content type, the standard does not say anything about visualisation. That means different server implementations will render a report differently. Nevertheless, the WADO service is a simple approach for accessing particular DICOM objects without requiring the client to "speak" DICOM. A number of commercial implementations supporting WADO is already available, which is not surprising, because standard can be implemented with little effort since many existing components like Web browsers, servers and DICOM viewers can be re-used. Although the standard does not specify where the unique identifiers of the DICOM objects are obtained, it provides a way to harmonise the HTTP query syntax of already existing DICOM enabled Web servers. Typical use cases for WADO are references to DICOM images from within a web-based EHR, references to DICOM images sent by e-mail or made available through a web browser for purposes of image and report distribution, both within or between healthcare enterprises. The anonymisation function in WADO is useful mainly for teaching and clinical studies.

5.2 DICOM Structured Reporting

In the year 2000, an extension to the DICOM standard has been officially released that covers medical reports and other clinical data [DICOM 2004]. Structured Reporting (SR) [CLUNIE 2000; Hussein et al. 2004a; 2004b] is a general model for encoding medical reports in a structured manner in DICOM's tag-based format. It allows the existing DICOM infrastructure network services like storage or query/retrieve, to be used to archive and to communicate, to encrypt and to digitally sign structured reports with only relatively small changes to existing systems. In addition to the header information that is also used for DICOM images, the actual content of a structured report is represented by a document tree. Each content item (node) of the tree contains some piece of information, for example, a text paragraph or a reference to an image. A set of well-defined relationships describe how "parent" and "child" content items in the hierarchical document structure are related to each other. The semantics of most content items in the SR document tree is described by a machine-readable code and, therefore, enables computer-supported automatic evaluation and processing.

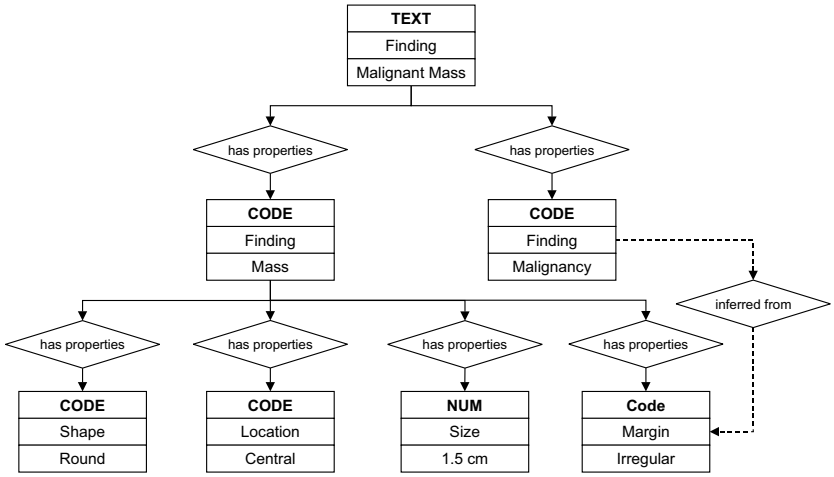

Fig. 7. DICOM SR Document Tree

Figure 7 shows an example of a DICOM SR document tree in which the malignancy of a mass is inferred from the observation that the mass has an irregular margin. The various terms and measurements can be encoded in a machine-readable way through annotation with codes taken from controlled vocabularies such as SNOMED [SNOMED] or LOINC [LOINC]. DICOM SR provides a very flexible model to store almost any kind of data ranging from simple free text reports to completely structured documents with numeric measurement values and codes. In order to enhance interoperability in practice, the DICOM standard specifies document classes and other types of constraints for different medical applications. For example, the standard defines templates to harmonise the document structure and groups of codes to limit the choice for a particular context. This collection of standard templates, context groups and codes is called DICOM Content Mapping Resource (DCMR).

Table II shows the definition of the DICOM SR template for a basic diagnostic imaging report (Template ID 2000 from [DICOM 2004, part 16]). Each line in the table corresponds to a single content item (node) or set of content items in an SR document. The nesting level (NL) column defines the tree structure and depth. The relationship with the parent content item in the tree and the value type (VT) of the node are defined and restrictions for the concept name, i. e. the machine readable code that defines the semantics attached to a node in the tree, are assigned by value or by reference to a table of possible values identified by its basic context group identifier (BCID). The value multiplicity (VM) defines if a tree node may appear only once or can be repeated. The requirement type (RT) defines whether the node is mandatory in each document or user optional. The condition column allows to describe conditions under which certain tree nodes are mandatory, allowed or forbidden. The value set constraint (VSC) field allows to define restrictions for the content of a tree node (as opposed to the concept name). The table also shows that DICOM templates make use of other templates,

Table II. DICOM SR Template for Basic Diagnostic Imaging Report [DICOM]

	NL	Rel with Parent	VT	Concept Name	VM	RT	Cond	VSC
1			container	BCID(7000) Diagnostic Imaging Report Document	1	M		Root node
2	>	has concept mod	code	EV(121058, DCM, "Procedure reported")	1-n	U		
3	>	has concept mod	include	DTID(1204) Language of Content Item and Descendants	1	M		
4	>	has concept mod	include	DTID(1210) Equivalent Meaning of Concept Name	1-n	U		
5	>	has obs context	include	DTID(1001) Observation Context	1	M		
6	>	contains	container	BCID(7001) Diagnostic Imaging Report Headings	1-n	U		
7	>>	has obs context	include	DTID(1001) Observation Context	1	U		
8	>>	contains	code	BCID(7002) Diagnostic Imaging Report Elements	1-n	U		
9	>>>	inferred from	include	DTID(2001) Basic Diagnostic Imaging Report Observations	1-n	U		
10	>>	contains	text	BCID(7002) Diagnostic Imaging Report Elements	1-n	U		
11	>>>	inferred from	include	DTID(2001) Basic Diagnostic Imaging Report Observations	1-n	U		
12	>>	contains	include	DTID(2001) Basic Diagnostic Imaging Report Observations	1-n	U		

which are expanded like macros and referred to by their DICOM template identifier (DTID). It should be noted, however, that the standard does not specify how an SR document is actually rendered by an application. The visualisation of reports for a human reader is regarded to be out of scope. Nevertheless, the application has to make sure that the full meaning of the report is conveyed in an unambiguous manner.

In the last couple of years DICOM Structured Reporting has been implemented many times but mainly as part of prototypes like the ones that vendors show at public IHE demonstrations. Only recently medical industry started to offer commercial products that include SR support. These products focus on medical fields that are already well-covered by the DICOM standard and where standard document templates exist, that is, radiology and cardiology. In these image acquiring departments, DICOM objects and services are already widely used and Structured Reporting is just the "missing link". Outside of the "imaging world" DICOM is not that common and, therefore, it is quite unlikely that Structured Reporting will

become accepted in the near future as an EHR standard.

6. ISO/TC 215 HEALTH INFORMATICS

The ISO technical committee on health informatics describes its scope as "standardisation in the field of information for health, and Health Information and Communications Technology (ICT) to achieve compatibility and interoperability between independent systems. Also, to ensure compatibility of data for comparative statistical purposes (e.g. classifications), and to reduce duplication of effort and redundancies." This scope includes the development of EHR standards. So far, however, ISO/TC 215 has not developed a complete EHR architecture but only provided individual building blocks. One such standard is "Web Access to DICOM Persistent Objects (ISO IS 17432)" which is presented in Section 5.1.

6.1 Interoperability and Compatibility in Messaging and Communication Standards (ISO TR 18307)

The ISO Technical Report TR 18307:2001 "Health informatics – Interoperability and compatibility in messaging and communication standards – Key characteristics" [ISOTR18307 2001] describes a set of key characteristics to achieve interoperability and compatibility in trusted health information interchange between application systems. In particular, "the interoperability needs of the healthcare community for the subject of care, the healthcare professional, the healthcare provider organisation, its business units and the integrated delivery network" are specified. The goal is to offer criteria for standards developers and implementers of standards for messaging and communications in the healthcare domain and to provide a guide for software developers and vendors, healthcare providers and end users. The technical report lists a number of fundamental principles and objectives to ensure trusted health information interchange. For example, it addresses issues like the completeness and accuracy of health records, the access rules and auditability, the data protection and integrity, and user authentication and accountability. From these principles the key characteristics are derived and further described.

6.2 Requirements for an EHR Architecture (ISO TS 18308)

The purpose of the ISO Technical Specification TS 18308:2004 "Health informatics – Requirements for an electronic health record architecture" [ISOTS18308] is "to assemble and collate a set of clinical and technical requirements for an electronic health record architecture (EHRA) that supports using, sharing, and exchanging electronic health records across different health sectors, different countries, and different models of healthcare delivery".

This specification only lists requirements for an electronic health record architecture but does not specify the architecture itself. Therefore, the primary target group are developers of EHR architecture standards like CEN EN 13606 (Section 3) and other reference architectures such as openEHR (Section 2). In fact, the first known compliance test against the ISO TS 18308 requirements has been done for the openEHR reference model.

6.3 EHR: Definition, Scope and Context (ISO TR 20514)

The ISO Committee Draft Technical Recommendation 20514 "EHR, definition, scope and context" [ISOTR20514 2004] describes a "pragmatic classification of electronic health records". Furthermore, it provides "simple definitions for the main categories of EHR and supporting descriptions of the characteristics of electronic health records and EHR systems".

This technical report makes a clear distinction between the content of the EHR and its structure. The so-called "basic-generic EHR" defines the structure of the EHR in a generic manner to ensure its applicability to a wide range of existing and future EHR users and systems. The definitions also support legislative and access control requirements that apply to all kinds of EHRs. The basic-generic EHR is supplemented by a more detailed and specialised definition to cover two of the most essential characteristics of the EHR: "the ability to share patient health information between authorised users of the EHR and the primary role of the EHR in supporting continuing, efficient and quality integrated healthcare". The principle definition of the EHR, which is a specialisation of the basic-generic EHR definition, is called the Integrated Care EHR (ICEHR).

7. INTEGRATING THE HEALTHCARE ENTERPRISE (IHE)

Integrating the Healthcare Enterprise (IHE) [Channin et al. 2001a; Channin et al. 2001b] is a non-for-profit initiative that was founded in 1998 in the USA by the Radiological Society of North America (RSNA) [RSNA] and the Healthcare Information and Management Systems Society (HIMSS)[HIMSS]. The initiative has meanwhile established working groups in Europe including national working groups in France, Germany, Italy, UK, Netherlands, Norway, Denmark, Spain and Sweden [Eichelberg et al. 2003] and also in Japan. The goal of the initiative as explained in the mission statement [IHEMS] is to stimulate integration of healthcare information resources.

While IHE does not develop standards as such, it selects and recommends appropriate standards for specific use cases and also develops restrictions, that is, application profiles for these standards that allow for a simplified system integration. The result of this technical work is published as the IHE Technical Framework [IHETF] and revised annually.

IHE is strongly supported by the industry: more than 110 companies have developed IHE compliant systems between 1999 and 2004 and participated in the cross-vendor testing events organised by IHE, including most of the market leaders in the modality, RIS and PACS sectors. This means that standards recommended by IHE have a high probability of a quick uptake in the medical market.

The IHE technical framework currently covers use cases which are called integration profiles for Radiology, Cardiology, Laboratory as well as "IT Integration" (IT-I). IT-I covers use cases for inter-departmental or inter-institutional system integration.

The following IT-I integration profiles are currently available:

—*Retrieve Information for Display (RID)* provides a simple and rapid read-only access to patient-centric clinical information that is located outside the user's current application. It supports access to existing persistent documents in well-

known presentation formats such as CDA, PDF and JPEG. It provides access to specific key patient-centric information such as allergies, current medications, summary of reports for presentation to a clinician.

—*Enterprise User Authentication (EUA)* is a means to establish one name per user that can then be used on all of the devices and software that participate in this integration profile. EUA greatly facilitates centralised user authentication management and provides users with the convenience and speed of a single sign-on, based on Kerberos [Kerberos] and the HL7 CCOW standard [HL7CCOW].

—*Patient Identifier Cross-referencing (PIX)* provides cross-referencing of patient identifiers from multiple Patient Identifier Domains. These patient identifiers can then be used by identity consumer systems to correlate information about a single patient from sources that know the patient by different identifiers.

—*Patient Synchronized Applications (PSA)* is a means for viewing data for a single patient using independent and unlinked applications on a user's workstation. It reduces the repetitive tasks of selecting the same patient in multiple applications. It is based on the HL7 CCOW standard [HL7CCOW].

—*Consistent Time (CT)*: This profile provides mechanisms such as Network Time Protocol (NTP) or Simple Network Time Protocol (SNTP) to synchronise the time base between multiple systems.

—*Audit Trail and Node Authentication (ATNA)* establishes security measures which, together with the Security Policy and Procedures of the enterprise, provide patient information confidentiality, data integrity and user accountability. Currently, this profile is in trial implementation.

—*Cross-Enterprise Document Sharing (XDS)* provides a standards-based specification for managing the sharing of electronic clinical documents with textual and structured content that healthcare enterprises (anywhere from a private physician to a clinic to an acute care in-patient facility) have decided to explicitly share. This contributes to the foundation of a shared Electronic Health Record. This profile is currently in trial implementation.

—*Patient Demographics Query (PDQ)* provides ways for multiple distributed applications to query a central patient information server for a list of patients, based on user-defined search criteria, and retrieve a patient's demographic (and, optionally, visit or visit-related) information directly into the application. This profile leverages HL7 2.5 and is currently in trial implementation.

—*Personnel White Pages (PWP)* provides access to basic human workforce user directory information, based on the Domain Name Service (DNS) and the Lightweight Directory Access Protocol (LDAP). Currently, this profile is in trial implementation.

Among these integration profiles, Retrieve Information for Display (RID) and Cross-Enterprise Document Sharing (XDS) both address how to access to electronic healthcare records in various formats and therefore they are further elaborated in the following sections.

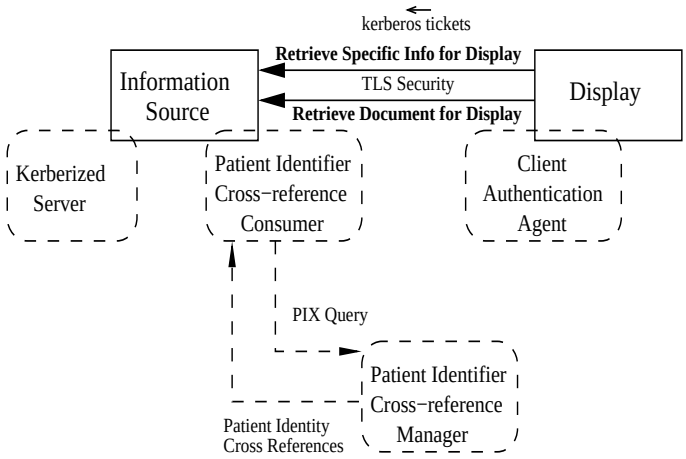

Fig. 8. Interaction between the RID, EUA, ATNA and PIX Integration Profiles

7.1 Retrieve Information for Display (RID)

IHE defined RID as a Web service by providing its WSDL (Web Service Description Language) [WSDL] description with a binding to HTTP GET. The WSDL description of different information sources may vary depending on how they implement the RID features, but the generic template provided by IHE very much restricts the possible variations.

For each integration profile, IHE defines "actors" which represent IT systems involved in the use case and "transactions" which define interfaces between actors. Each transaction is based on an existing interface standard. In the case of RID, there are only two actors: The "information source" is a system that maintains a database of persistent clinical documents and specific key patient-centric information such as allergies, current medications, and summary of reports. The "display" is a system that accesses the information source, retrieves patient-centric information or persistent documents, and displays them to a human observer.

The focus of the integration is visual presentation, not a complete integration of the structured databases on which the actors might be based. Documents are exchanged in well-known presentation formats such as HL7 CDA Level One, PDF or JPEG. Note that HL7 CDA Level One provides little more structure than a web page and it can easily be converted into a XHTML page using XSLT transformations and appropriate style sheets. It is the responsibility of the information source to convert the healthcare specific semantics into a suitable presentation format. The display, on the other hand, may process and render this presentation format with only generic healthcare semantics knowledge, but will in general not be able to provide any processing of the healthcare information beyond document display.

The RID integration profile does not address access control or security of information in transit. This can be implemented using the EUA and ATNA profiles. The IHE technical framework explicitly states that by linking (RID) with two other IHE profiles, EUA and PIX, this profile's reach can extend across organ-

isation boundaries within an enterprise. In this scenario, the information source would be grouped with a "patient identifier cross-reference consumer" actor from the PIX profile. This would enable the information source to serve requests to patient IDs from a different patient ID domain by looking up a cross reference in the PIX "patient ID cross-reference manager". When combined with EUA, the information source would act as a "kerberized server", that is, validate the Kerberos tickets provided as part of the HTTP requests. The display actor would be grouped with a "client authentication agent" and provide Kerberos tickets together with the requests. Finally, communication could be protected using the Transport Layer Security (TLS) protocol encapsulating the HTTP traffic, as described in the ATNA integration profile. The Interaction among RID, EUA, ATNA and PIX Integration Profiles are depicted in Figure 8.

The RID Integration Profile is not mainly intended for cross-enterprise document exchange since the display actor needs a-priori knowledge about the patient ID of the patient to query for. In the case of a hospital with a single HIS system, this could be an HL7 ADT feed (see Section 4) providing all systems with notification messages about all patient admissions, discharges and updates to demographics. In the case of hospitals using multiple HIS systems (i. e. multiple different patient ID domains), RID is still applicable if a cross-index of patient IDs such as a master patient index exists. In this case RID needs to be combined with PIX as described above. In the more general case, however, if two hospitals that do not share patient IDs and do not maintain a master patient index, RID is simply not applicable because there would be no way for the display to query anything from the server. Other than that, there is no obvious reason why RID should not be used for cross-enterprise document access.

The RID profile was initially published in August, 2003 and has seen a rather quick market uptake. Prototype implementations from 22 different vendors have been successfully tested for their interoperability at the IHE cross-vendor testing events between 2003 and 2005, and several commercial products are already available on the market.

7.2 Cross-Enterprise Document Sharing (XDS)

The basic idea of IHE XDS is to store healthcare documents in an ebXML registry/repository to facilitate their sharing. IHE XDS is not concerned with document content; it only specifies metadata to facilitate the discovery of documents.

In the IHE XDS Profile, a group of healthcare enterprises that agree to work together for clinical document sharing is called the "Clinical Affinity Domain". Such institutes agree on a common set of policies such as how the patients are identified, the access is controlled, and the common set of coding terms to represent the metadata of the documents.

As already mentioned, IHE XDS handles healthcare documents in a content neutral way, that is, a document may include any type of information in any standard format such as simple text, formatted text (e.g., HL7 CDA Release 1), images (e.g., DICOM) or structured and vocabulary coded clinical information (e.g., CDA Release 2, CEN ENV 13606 or DICOM SR). Given this, to ensure the interoperability between the document sources and the document consumers, the clinical affinity domains also agree on the document format, the structure and the content.

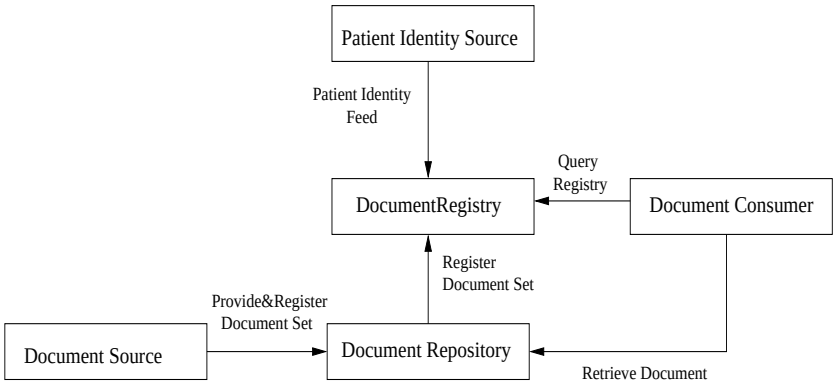

Fig. 9. Actors and Transactions of the XDS Integration Profile [IHE-XDS 2004]

The XDS Integration Profile is not intended to address all cross-enterprise EHR communication needs. Some scenarios may require the use of other IHE Integration profiles, such as Patient Identifier Cross-referencing, Audit Trail and Node Authentication, Enterprise User Authentication, and Retrieve Information for Display. Other scenarios may be only partially supported, while still others may require future IHE Integration profiles, which will be defined by IHE as soon as the necessary base standards are available.

The actors and transactions defined by the XDS profile are depicted in Figure 9. The three systems in the centre of the figure provide the document registry that is used to store and interchange clinical documents that together comprise the "longitudinal", that is, life-long, cross-institutional healthcare record, dubbed EHR-LR (EHR – Longitudinal Record) in IHE terminology. The document source and document consumer actors would be typical clinical systems used at the point of care. Information stored in these systems are part of the care-delivery record, dubbed EHR-CR (EHR – Care-delivery Record) in IHE terminology. Once an episode of care is completed, clinical information is converted into document format if necessary and delivered to the EHR-LR which is patient centric as opposed to hospital centric. The XDS actors are as follows (Figure 9):

—*Document Source* represents a healthcare point of service system where care is provided and associated clinical information is collected. A document source provides clinical documents and metadata to one of the EHR-LR document repositories which in turn forwards the metadata to the central document registry.

—*Document Registry* provides handling of the metadata for all published clinical documents that may be queried. The document registry does not actually store any document, it just stores meta-information along with references to the document repository where the document can be retrieved. The document registry is basically an ebXML registry, along with some IHE-specific extensions such as the "XDS registry adaptor function" which validates incoming requests prior to submitting them to the ebXML registry and guarantees the atomicity of submission operations by maintaining certain status flags in the registry.

—*Document Repository* maintains and stores published documents that may be

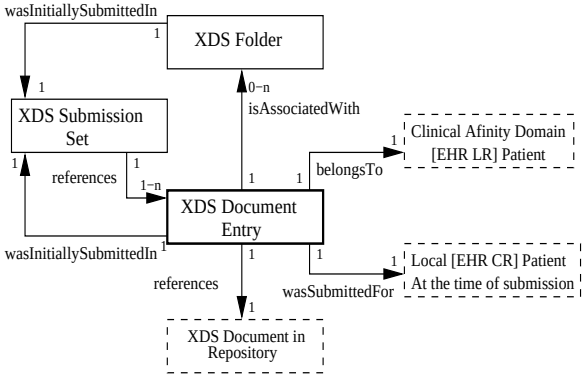

Fig. 10. Information Model of the XDS Integration Profile [IHE-XDS 2004]

retrieved. A document repository cannot be queried. Document consumers need to know a document unique identifier in order to request a document from the repository.

—*Document Consumer* is a service application system where care is provided and access to clinical documents is needed.

—*Patient Identity Source* is the central system that assigns and manages patient identifiers for the XDS affinity domain.

As mentioned previously, the XDS design is document centric, that is, a document is the smallest unit of information provided in a document registry. IHE defines an XDS Document as "a set of attested clinical information (structured or not) which form an element of a patient record to be shared." Each document is assigned to a single patient. In addition to the patient reference, there are two concepts that can be used to structure the set of documents available in a document registry as presented in Figure 10:

—An XDS Submission Set is defined as "a set of documents related to a patient that a (team of) clinician(s) in the same source system have decided to make available to potential consumers." A document source always provides and registers a submission set as part of a "Provide and Register Document Set" transaction. The processing of the document set is atomic, i. e. all documents of the set are registered in the repository and the registry, or the complete operation is rejected.

—An XDS Folder is a logical grouping of documents belonging to a single patient. Each document source can create labelled folders and assign documents to folders. The folders are virtual in the sense that one document can be contained in zero, one or multiple folders.

The XDS integration profile defines five transactions between the actors involved (Figure 9). Together these allow both a read access and a write access to the XDS registry and repositories:

—*Patient Identity Feed* is a transaction adapted from the Patient Identifier Cross-referencing (PIX) integration profile. Its purpose in the XDS Integration Profile is to populate the registry with patient identifiers that have been registered for the affinity domain. This transaction is based on HL7 version 2.3.1 and uses the ADT message (admission, discharge, transfer) with various HL7 trigger events for patient registration, update and merge. Patient demographics have to be communicated through this transaction before any document for the corresponding patient can be registered in the document registry.
It should be noted that a standard ebXML registry cannot receive and process ADT messages. Therefore IHE has defined an XDS "Registry Adaptor" which is a set of functionality that is not provided by the ebXML registry standard, but is instead specified by XDS to support integration into the healthcare environment. For example, it is the XDS "Registry Adaptor" which processes ADT messages. Further information on XDS "Registry Adaptor" is provided in Section 7.3.

—*Provide and Register Document Set* is the only transaction defined for the document source actor. It is used to submit an XDS submission set to a document repository. The repository is responsible to persistently store these documents, and to register them in the document registry using the "Register Documents" transaction by forwarding the document metadata received from the document source. A submission set may contain documents (which are included as byte streams), the metadata needed to correctly register the documents, folders to be created and assignments of documents to folders.
The message submitted uses the ebXML messaging framework, which is SOAP with MIME attachments, the first attachment being the ebXML SubmitObjectRequest message, followed by a separate MIME attachment for each document submitted. XDS exactly defines the set of metadata to be submitted for a document, folder or document set, based on the ebXML Registry Service (ebRS) [ebRS] and information model (ebRIM) [ebRIM] as detailed in Section 7.5.

—*Provide and Register Document Set transaction* stores the submitted documents and generates a URI for each document. It then initiates a "Register Document Set" transaction which forwards the ebXML metadata received from the document source along with the document URIs to the central document registry. The IHE technical framework states that the Document Registry Actor ensures that document metadata is valid before allowing documents to be registered. If one or more documents fail the metadata validation, the Register Document Set transaction fails as a whole. In the document registry, the "XDS Registry Adaptor" maintains the overall consistency of the registry.

—*Query Documents* is a transaction issued by the document consumer. It returns a list of document entries that contain metadata found to meet the specified criteria including the locations and identifier of each corresponding document in one or more Document Repositories. The query uses the SQL language as specified by the ebXML Registry Services.

—*Retrieve Document* is a transaction initiated by a document consumer to retrieve a document with known HTTP URI from a document registry. The registry simply returns the requested document in the HTTP response.

XDS defines a certain lifecycle for all documents registered in the document registry.

Initially, each document is registered as "submitted" in the registry. As soon as the Registry Adaptor (Section 7.3) has processed all documents and folders in one submission set and has decided that the atomic submission can be accepted, the status is updated to "approved" and the documents are available for query and retrieval. Under the responsibility of the original submitter, documents may be declared "deprecated" later, that is, the documents may become obsolete but still available for query and retrieval.

In addition to the document status, document relationships can be defined. A document may be defined to be a "replacement" for another document in the repository, which would then be marked as deprecated. A document may be defined to be an "addendum" for another document in the repository. Finally, a document may be defined to be a "transformed" version of another document. This relationship is used to describe machine translation between different formats, e. g. a PDF file created from a CDA original or a DICOM Structured Report. Typically both versions of the document would then remain approved in the repository.

Different implementation strategies of the XDS integration profile are possible. One model would integrate the document source and the document repository, that is, each participating organisation would maintain a repository of its own documents and only register the documents in a central registry available to everyone in the affinity domain. Another model would set up a third-party repository, that is, there would be one or more central repositories to which all participating document sources would submit documents. In this case, the EHR-LR would necessarily be a copy of the local EHR-CR records, thus possibly increasing the total storage volume. The choice of implementation strategy certainly depends on the use case, that is, on the question which organisation would be responsible to maintain which actor or set of actors within one affinity domain, which could be corporate, community, regional or even national.

Even though the XDS profile is rather new (the "trial implementation draft" has been published in August, 2004), XDS prototype implementations from 28 distinct vendors have been successfully tested for their interoperability at the IHE cross-vendor testing events, including 5 different registries and 12 different repositories. The first commercial products are also available already. This indicates a significant interest from industry and indicates a quick market uptake to be very likely.

7.3 XDS Registry Adaptor

There is certain functionality required by the IHE XDS specification that is not supported by ebXML registry specification. For example, XDS requires the submitted metadata to be validated. Such functionality that is not available in the ebXML registry standard, is provided by the XDS "Registry Adaptor".

The XDS "Registry Adaptor" can be described as a pre-processor for the ebXML registry, maintaining the overall consistency of the registry. It validates the patient ID, document MIME type, metadata, coded values and document classification entries in the submission. It verifies that all documents in a folder belong to the same patient and maintains a "last update time" attribute for each folder. It manages document amendments and replacements. Finally it ensures that submission of multiple documents is an atomic operation.

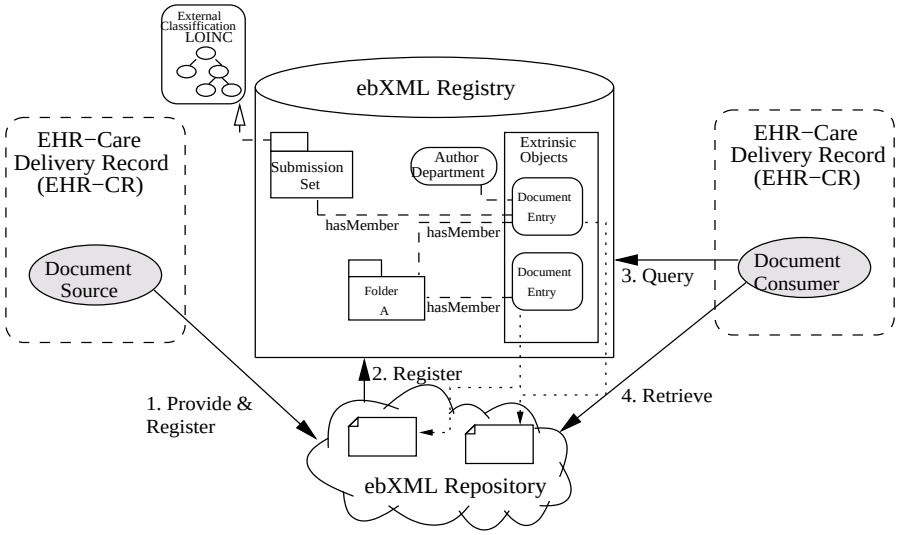

Fig. 11. IHE XDS Technical Architecture

7.4 IHE XDS ebXML Technical Architecture

In this section, we describe how various constructs of IHE XDS are actually implemented in the ebXML framework. As shown in Figure 11, an XDS document, stored in the repository, is represented as an "ebXML ExstrinsicObject" in the registry. "ExtrinsicObjects" in ebXML are used to provide metadata that describes submitted content whose type is not intrinsically known to the registry. As already mentioned, in order to group the related documents, the XDS documents are organized into folders (e.g. a period of care, a problem, immunizations, etc.). In this way a document consumer can find all the entries placed in the same folder. XDS document folders are constructed in the ebXML registry by using "ebXML RegistryPackage". As previously mentioned ebXML RegistryPackage is used to group logically related "RegistryObject" instances together. Folders cannot be nested since registry packages cannot be nested in ebXML.

In order to support atomic submission of documents to the registry and to make a permanent record in the registry of objects, XDS documents are submitted as "XDS SubmissionSet"s. The submission sets are also represented through "ebXML RegistryPackage" constructs and submitted by using "ebXML SubmitObjectsRequest". An "XDS Submission Request" includes information on the metadata of the documents as well as the folders that the documents belong and/or new folders to be created.

7.5 IHE XDS Metadata

XDS specifies a set of predefined metadata elements to be associated with XDS documents, submission sets and folders to facilitate their discovery. Some of the metadata elements are straightforward document properties such as "authorDepartment" or "creationTime". Healthcare domain specific semantics is provided

Table III. Selected XDS Document Metadata Attribute Definitions

XDS Document Attribute	*Definition*	*ebRIM Attribute Type*
authorDepartment	Represents a specific department within a healthcare facility under which the human and/or machines authored the document.	Slot
classCode	The code specifying the particular kind of document (e.g. Prescription, Discharge Summary, Report). It is suggested that the XDS Affinity Domain draws these values from a coding scheme providing a coarse level of granularity	External Classification
classCodeDisplay-Name	The name to be displayed for communicating to a human the meaning of the classCode	Name attribute on Classification object, coding scheme name saved as codingScheme slot on Classification
parentDocument Relationship	The type of relationship that the document has with the parentDocument (e.g. Replace, addendum, or transformation).	Association Type
eventCodeList	This list of codes represents the main clinical acts, such as a colonoscopy or an appendectomy, being documented	External Classification
practiceSettingCode	The code specifying the clinical specialty where the act that resulted in the document was performed	External Classification
typeCode	The code specifying the precise kind of document (e.g. Pulmonary History and Physical, Discharge Summary, Ultrasound Report)	External Classification

only through "classCode"s. A "classCode" is represented as an *External Classification* in ebXML registry. There are a number of native *External Classifications* in the ebXML registry implementation (OMAR [FREEebXML]) such as North American Industrial Classification Scheme (NAICS) [NAICS] and Universal Standard Products and Services Classification (UNSPSC) [UNSPSC]. Other types of external classifications have be to introduced to the registry explicitly. As an example, Figure 12 shows how the "Logical Observation Identifiers Names and Codes (LOINC)" [LOINC] External Classification denoting laboratory codes can be introduced to the ebXML registry. After introducing such an External Classification to the registry, in order to use it, its unique identifier, called UUID which is assigned by the registry is necessary. The query shown in Figure 13 is used to retrieve such a UUID.

Table III gives some example metadata attributes for XDS Documents such as "authorDepartment", "classCode" "classCodeDisplayName" and "parentDocument Relationship". Such metadata is defined in the registry through ebXML RIM semantic constructs. For example, "authorDepartment" is defined as a slot of the ExtrinsicObject representing an XDS Document in the registry. "parentDocument Relationship", on the other hand is defined as an "Association Type" which extends the predefined Association types of ebXML registry with 3 new values,

```
<SubmitObjectsRequest>
 <rim:RegistryObjectList>
  <rim:ClassificationScheme id="loinc" isInternal="false"
   nodeType="urn:uuid:bd0c092a-cb38-4578-8520-81274054f678">
     <rim:Name>
      <rim:LocalizedString charset="UTF-8" value="LOINC"/>
     </rim:Name>
     <rim:Description>
      <rim:LocalizedString charset="UTF-8" value="LOINC coding"/>
     </rim:Description>
  </rim:ClassificationScheme>
 </rim:RegistryObjectList>
</SubmitObjectsRequest>
```

Fig. 12. A SubmitObjectsRequest introducing the LOINC Term List to the ebXML registry

namely, "append" (APND), "replace" (RPLC) and "confirm" (XFRM).

```
<AdhocQueryRequest >
 <query:ResponseOption returnComposedObjects="true"
  returnType="LeafClassWithRepositoryItem"/>
 <rim:AdhocQuery id="tempId">
  <rim:QueryExpression queryLanguage=
   "urn:uuid:c26215e8-7732-4c7f-8b04-bd8115c325e9">
    SELECT * FROM CLASSSCHEME WHERE ID IN
      (SELECT PARENT FROM NAME WHERE VALUE LIKE 'LOINC')
  </rim:QueryExpression>
 </rim:AdhocQuery>
</AdhocQueryRequest>
```

Fig. 13. A Query to find the UUID of the External Classification submitted in Figure 12

Similarly, the metadata for submission sets, such as "authorDepartment" is defined as the slot of the "ebXML RegistryPackage" object. On the other hand, "contentTypeCode" values are drawn from a vocabulary defined by the affinity domain and use External Classification. Here, a submission set is related with a node in the External Classification by using "ebXML Classification". As an example, "MyXDSSubmissionSet" is related with "LOINC's Hospital Discharge Summary Note" as shown in Figure 14 by classifying "MyXDSSubmissionSet" through a Classification node whose nodeRepresentation is "34105-7" (corresponding to LOINC's Hospital Discharge Summary Note code) and whose classificationScheme value is "urn:uuid:c82d4b0b-d042-4d49-a5d1-8ee5153aea9f" which is the UUID obtained as a result of the query shown in Figure 13.

7.6 Querying the Registry

The metadata defined for XDS documents, folders and submission sets can be used to retrieve the desired documents by querying the registry through SQL. Example queries include retrieving specific types of documents of a patient for a time interval, or by author person, or all documents in a folder or a submission set.

All queries return either metadata for one or more registry objects, or object references for one or more registry objects (registry UUIDs). To query the registry, the user has to know the underlying relational schemas of the ebXML registry specification.

```
<SubmitObjectsRequest >
 <rim:RegistryObjectList>
  <rim:RegistryPackage id = 'MyXDSSubmissionSet'>
   <rim:Name> <rim:LocalizedString value = 'Example '/>
   </rim:Name> </rim:RegistryPackage>
  <rim:Association id = 'hasMember' associationType =
  'urn:uuid:2d03bffb-f426-4830-8413-bab8537a995b'
  sourceObject = 'MyXDSSubmissionSet'
  targetObject = 'LabResuts'/>
  <rim:ExtrinsicObject id="LabResults" mimeType="text/xml">
    <rim:Name> <rim:LocalizedString value = 'ExampleLR'/>
    </rim:Name>
  </rim:ExtrinsicObject>
  <rim:Classification nodeRepresentation="34105-7"
    classifiedObject="MyXDSSubmissionSet"
    id="LOINC's Hospital Discharge Summary Note"
    classificationScheme=
    "urn:uuid:c82d4b0b-d042-4d49-a5d1-8ee5153aea9f" />
 </rim:RegistryObjectList>
</SubmitObjectsRequest>
```

Fig. 14. An Example to classifying a Submission Set with an External Classification

Consider for example the query given in Figure 15 which retrieves all approved "Hospital Discharge Summary Notes" of a patient whose patient id is '123'. The schemas of the relations accessed in this query are given in Table IV. In this query, the External Classification, namely, Logical Observation Identifiers Names and Codes (LOINC) [LOINC] is used which was introduced to the registry as shown in Figure 12. The "classificationScheme" value is "urn:uuid:c82d4b0b-d042-4d49-a5d1-8ee5153aea9f" which is the UUID obtained as a result of the query shown in Figure 13. The value of "nodeRepresentation" attribute is 34105-7 from LOINC which denotes "Hospital Discharge Summary Note".

Table IV. ebXML Relational Schemas

ExtrinsicObject(id, home, objectType, status, expiration, majorVersion, minorVersion, stability, userVersion, isOpaque, mimeType)
ExternalIdentifier(id, home, objectType, status, registryObject, identificationScheme, value)
Classification(id, home, objectType, status, classificationNode, classificationScheme, classifiedObject, nodeRepresentation)

8. MEDICAL MARKUP LANGUAGE (MML)

The Medical Markup Language (MML) [Araki et al. 2000; Guo et al. 2004] has been developed since the mid 1990s by the Electronic Health Record Research Group of the Japanese Ministry of Health and Welfare. Its purpose is to provide a standardised way to exchange medical documents and other clinical data. The first version of this specification was based on SGML (Standard Generalized Markup Language) but later XML (Extensible Markup Language) was chosen. The current version 3.0 [MML] uses the XML-based HL7 Clinical Document Architecture Release One (CDA) format (Section 4.1) with a local header extension to store MML specific header fields and local markup to store the MML specific content.

```
SELECT * FROM ExtrinsicObject doc
WHERE
doc.id in (SELECT doc.id FROM ExtrinsicObject
doc, ExternalIdentifier ei
WHERE
  doc.objectType=XDSDocumentEntry AND
  ei.identificationScheme=XDSPatientId AND
  ei.registryObject=doc.id AND
  ei.value='123') AND
  doc.id in
  (SELECT classifiedObject FROM Classification
    WHERE classificationScheme=
    'urn:uuid:c82d4b0b-d042-4d49-a5d1-8ee5153aea9f' AND
    nodeRepresentation='34105-7') AND
    doc.Status='Approved'
```

Fig. 15. An Example XDS SQL Query to retrieve all discharge summaries for a patient

MML documents can be exchanged via HL7 messages or by any other means of electronic communication. The local header contains many fields that are also stored in the CDA header (e. g. patient demographics, document creator and diagnostic information) but in a different format. Even the content of the document is duplicated to some extent in the local markup section of the document body. So currently, HL7 CDA is merely used as a standardised container that carries MML information. However, compared to CDA Release One, MML specifies more restrictions on the structure and the content of a document. The Medical Markup Language was never really used outside Japan and with the appearance of HL7 CDA there seems to be no reason that this will change in the future. However, within Japan, MML seems to be actively used and there are commercial products on the market that support MML.

9. ANALYSIS OF EHR STANDARDS

Because of the convergence of CEN EN 13606 (EHRcom) and GEHR/openEHR into a harmonized EHR architecture (details on this process are provided in Section 9.5), the latter is not examined separately in the following discussion, but is subsumed under the term EHRcom. The remaining seven EHR related standards that are surveyed vary widely regarding their scope and content. Some of them, for example CDA and MML only specify a content format but no communication protocol, others, for example XDS, only specify communication protocols and are "content agnostic", that is, they do not define any content format. Table V summarizes the scope of the EHR standards with regard to two basic properties of an EHR: the EHR content structure and the access services for locating, retrieving and submitting EHR content.

The only EHR standards that specify both content structure and access services are DICOM SR and EHRcom, although the communication protocol for EHRcom (EN 13606-5) is not yet published. WADO is somewhat of a special case because it only provides an alternative access protocol to DICOM images and reports, that is, it uses DICOM SR as a format for structured EHR content.

9.1 Analysis of the EHR Content Structure

Table VI summarises the functionality of the EHR standards regarding content structure. WADO, RID and XDS are not shown in the table since they do not

Table V. Scopes of EHR Standards

	CEN EHRcom	HL7 CDA	ISO WADO	DICOM SR	IHE RID	IHE XDS	MML
EHR Content Structure	Yes	Yes	No	Yes	No	No	Yes
EHR Access Services	Yes	No	Yes	Yes	Yes	Yes	No

define a content structure of their own.

Table VI. Analysis of EHR Standards' Content Structure

	CEN EHRcom	HL7 CDA	DICOM SR	MML
EHR contains persistent documents	Yes	Yes	Yes	Yes
EHR can contain multimedia data (images, signals, movies)	Yes	Yes	Yes	Yes
EHR document can contain references to multimedia data	Yes	Yes	Yes	Yes
EHR structured content suitable for processing	Yes	Yes	Yes	Yes
EHR supports archetypes / templates	Yes	Yes	Yes	Yes
EHR specifies library of archetypes / templates	Yes	Yes	Yes	Yes
EHR specifies distribution rules	Yes	No	No	Yes
EHR standard covers visualisation	No	Yes	No	No
EHR supports digital signatures on persistent documents	No	No	Yes	No

The table shows that the properties of the four content standards are remarkably similar. All of them can be used to store persistent structured documents. In the conventional (non-digital) medical workflow, the exchange of medical information between healthcare professionals is most usually organised in the form of documents. The aggregation of basic units of information such as patient demographics, individual measurements, procedure reports, diagnostic information and recommendations into complex documents such as discharge letters or diagnostic reports is not only a requirement of conventional mail transport; it also acknowledges the medico-legal requirement that somebody has to take over legal responsibility for a clinical document, usually in the form of a signature.

In contrast, medical information systems usually store medical information in a structured, normalised database in which each basic unit of information is kept separately. This allows to precisely retain the semantics of each individual entry in the patient record. However, medical information systems still allow to combine such entries into a classical document which might be exchanged in digital form between departments, for example by using HL7 messages or in the printed form. For CDA, DICOM SR and MML, the persistent document is the basic unit in which information is assembled, stored and communicated. In EHRcom, communication is based on the "EHR extract", which is more than a document because it contains one or

more "compositions" (which are roughly equivalent to documents) plus structuring information such as folders. However, content submission and audit trail are based on the composition, that is, persistent documents can be represented in EHRcom as well.

All of the EHR standards allow to add multimedia content (images, signals and movies) to the healthcare record, and all of them allow to reference multimedia data from within the structured content. In all cases, the EHR standards enable (but do not require) a fine-grained machine readable representation of clinical data using tree or graph structure and controlled vocabularies.

All of the content standards support the concept of a two-level modelling of EHR content using a simple reference model and an additional set of constraint rules that describe how certain clinical observations can be expressed in an unambiguous manner using structures of the basic reference model. In EHRcom, the constraint rules are called "archetypes". HL7 CDA has the equivalent concept of "templates", which are largely undefined at the moment, but are intended to be compatible with EHRcom archetypes through a common archetype definition language. In DICOM, the constraint rules are also called "templates" and their expressive power are very similar to that of EHRcom archetypes. MML "content modules" are less expressive as they reflect what can be specified with a conventional DTD (Document Type Definition). All of the standards aim at specifying a library of standard archetypes or templates, but the current status is quite different. DICOM SR already has an established set of standard templates that have been developed over years. MML specifies a small standard library of about a dozen DTD fragments, while EHRcom and CDA are still at the very beginning.

Two of the EHR content standards, namely EHRcom and MML, allow to specify "distribution rules", that is, statements specifying under which circumstances the EHR content may be communicated. Distribution rules are a different concept from access control in that they are part of the EHR content, not part of an EHR access protocol. It is up to the implementation of an EHR system how the distribution rules are executed or mapped to an EHR access protocol.

It is an interesting fact that most EHR standards do not specify how EHR content should be visualised, because this touches safety issues. The CDA specification gives at least a few recommendations on how documents are to be structured and encoded in order to facilitate the visualisation process. However, this requires that structured and coded content has to be stored redundantly in the document.

A final property in which the standards differ is the ability to attach digital signatures to EHR documents. DICOM SR is the only standard in which this is explicitly specified, through an extensive set of rules which address some peculiarities of the binary encoding used in DICOM. EHRcom, CDA and MML can all be represented in XML, so digital signatures could be handled through the XML-Signature standard. However, none of these standards explicitly states how signed documents would be handled by a document repository (i. e. EHR system).

9.2 Analysis of EHR Access Services

Table VII summarises the functionality of the EHR standards regarding access services. In this table, CEN EHRcom refers to the EHRcom communication protocol to be published as EN 13606-5 and DICOM SR refers to the DICOM Storage and

Query/Retrieve Services that are used for the transport of DICOM SR documents and images. CDA and MML which do not currently define access services are not shown in the table.

Table VII. Analysis of EHR Standards' Access Services

	CEN EHRcom	ISO WADO	DICOM SR	IHE RID	IHE XDS
Service for querying EHR content	Yes	No	Yes	Yes	Yes
Service for retrieving EHR content	Yes	Yes	Yes	Yes	Yes
Service for submitting EHR content	Yes	No	Yes	No	Yes
Document-centric storage / retrieval	No	Yes	Yes	Yes	Yes
Content format agnostic	No	No	No	Yes	Yes

The first three rows in Table VII show the support for the three basic EHR access services in the EHR standards: content query, retrieval and submission. WADO is a special case since it is the only protocol that does not allow to query for document content. Essentially, the requestor needs to know a priori the location and the document ID (Study, Series and Instance Unique Identifier) of a DICOM report or image in order to be able to retrieve the document. This shows that WADO is not a comprehensive EHR protocol of its own, but only a supplementary service that needs to be combined with other services such as XDS for queries. All standards support the retrieval of EHR content, but WADO and RID are "read only" services, that is, they do not support the submission of new documents to a repository.

In all standards, except EHRcom, the persistent document is the basic unit of information that is queried, retrieved or submitted to the EHR. EHRcom instead uses the concept of an "EHR extract", which may contain several documents, for query and retrieval (but not necessarily for submission).

Two of the EHR protocols, IHE RID and IHE XDS are content format agnostic, i. e. they treat documents as an opaque byte stream and only process the metadata accompanying the document. This approach has both advantages and disadvantages:

—*Advantages*: A content format agnostic EHR protocol is certainly significantly easier to define and implement than a comprehensive EHR architecture that covers both content and the service level. It also facilitates the inclusion of generic (non-medical) document formats such as PDF, RTF, JPEG or MPEG into the EHR. This is important because such generic formats are still intensively used in most healthcare enterprises.

—*Disadvantages*: Since the content format agnostic EHR may contain "anything", the EHR does not contain any structure beyond the structure established through the metadata in the EHR registry. It is in the general case not possible to support advanced services beyond document visualisation such as document processing, mediation, or automated translation services. It is not guaranteed that references between documents in the EHR are possible due to the varying content

formats. Finally, it is not guaranteed that a recipient will even be able to correctly visualise a document retrieved from the EHR system due to an unknown file format or encoding. RID works around this problem by requiring the document server to be able to convert each document into one of a few well-known formats (JPEG, CDA or PDF). XDS uses the concept of so-called XDS Document Content Profiles to address this problem. A document content profile restricts document formats and encoding options for specific clinical applications. An actor that claims conformance to a specific document content profile will be able to interoperate with other actors supporting the same profile. Most XDS document content profiles are under developement currently, but the first two profiles have been published. The Document Digital Signature content profile [IHE-DDS 2005] describes how digital signatures can be handled in an IHE XDS affinity domain, and the Cross-enterprise Document Sharing for Imaging (XDS-I) content profile [IHE-XDS-I 2005] describes how DICOM images can be archived and retrieved in XDS. This profile is based on the idea that instead of registering every single DICOM image as an XDS document, only a summary document in the form of a so-called "DICOM Key Object Selection Document" is submitted to XDS, and this document contains references that can be used by the document consumer to retrieve all related images using either WADO or the conventional DICOM network services.

9.3 Analysis of Security Features of EHR Standards

Table VIII shows the support for IT security features in the EHR protocol, as far as they are specified in the standards. All of these protocols can certainly be extended with security features for encryption, user credentials and access control, but such security concepts will only be interoperable among different implementations if standardised. Since the EHRcom protocol is not specified yet, its security properties are not known.

Table VIII. Analysis of Security Features of EHR Standards

	ISO WADO	DICOM SR	IHE RID	IHE XDS
Supports transport level encryption	Yes	Yes	Yes	Yes
Protocol allows to transmit user credentials	Yes	Yes	Yes	Yes
Protocol enforces access rules	No	No	No	Yes

All remaining four protocols support a transport level encryption, that is, an encryption of record content during transmission. In all cases this is implemented using the TLS (Transport Layer Security) protocol specified in [Dierks and Allen 1999]. All four protocols also optionally allow to transmit user credentials from the requesting system to the EHR system. This is an important feature because it allows the EHR system to determine which user is requesting access and, based on this information, to implement user or role based access control.

In the case of WADO and RID this can be implemented using the IHE Enterprise User Authentication Integration Profile (EUA) which is based on Kerberos

[Kerberos] and the transmission of Kerberos tickets through the HTTP protocol that is used both in WADO and RID.

In the case of DICOM SR, support for the communication of user credentials through the DICOM network protocol has recently been added to the standard [DICOM Supplement 99 2005] and this will also allow for the use of IHE EUA in DICOM. IHE is developing a new integration profile named Cross-Enterprise User Authentication (XUA) [IHE-XUA 2005] that addresses the issue of user authentication for cross-enterprise communication, in particular for use with the RID and XDS integration profiles. XUA is based on the OASIS Security Assertion Markup Language 2.0 (SAML) [SAML Overview 2005; SAML Profiles 2005]. Finally, IHE XDS is the only EHR protocol into which the enforcement of access rules will be built-in (as opposed to the mere provision of user credentials as a "hook" for a proprietary implementation of access control), but this functionality is only announced for the future and technical details are not known at this time.

9.4 Combining Different EHR Standards

Since some EHR standards only specify content structure and others only specify access services, it makes sense to consider the combination of EHR content structure and EHR access services from different standards. Table IX shows how the five EHR access protocols can be combined with the four EHR content formats. The table heading shows the communication protocols that are used by the EHR access services: EN 13606-5 for EHRcom which is yet to be specified; a HTTP URL formed with certain rules for WADO; the DICOM Storage and Query/Retrieve Service Class for DICOM SR; a set of web services (WDSL and HTTP GET) for RID and ebXML which in turn uses SOAP with a default binding to HTTP for XDS.

Table IX. Possible Combinations of EHR Content and Communication Protocol Standards – EHR Standard Content Formats

Protocol/ Content	EHRcom EN 13606-5	ISO WADO HTTP URL	SR DICOM Q/R	RID WSDL/HTTP	IHE XDS ebXML+HTTP
EHRcom extract	Yes	No	No	Yes	Yes
CDA document	No	No	No	Yes	Yes
SR document	No	Yes	Yes	Yes	Yes
MML document	No	No	No	Yes	Yes

In addition to the EHR content formats, there are a number of general purpose content formats that may be used in exchanging EHRs. Table X, shows how such formats can be combined with EHR access protocols.

Finally Table XI depicts how a number of multimedia content formats can be combined with EHR access protocols.

Table X. Possible Combinations of EHR Content and Communication Protocol Standards – General Purpose Content Formats

Protocol/ Content	EHRcom EN 13606-5	ISO WADO HTTP URL	SR DICOM Q/R	IHE RID WSDL/HTTP	IHE XDS ebXML + HTTP
PDF	No	Yes	Yes	Yes	Yes
HTML/ XHTML	No	Yes	No	Yes	Yes
RTF	No	Yes	No	Yes	Yes

Table XI. Possible Combinations of EHR Content and Communication Protocol Standards – Multimedia Content Formats for Images, Signals and Movies

Protocol/ Content	EHRcom EN 13606-5	ISO WADO HTTP URL	SR DICOM Q/R	IHE RID WSDL/HTTP	IHE XDS ebXML + HTTP
DICOM	No	Yes	Yes	Yes	Yes
JPEG	No	Yes	Partially	Yes	Yes
TIFF, BMP, PNG	No	No	No	Yes	Yes
MPEG	No	No	Partially	Yes	Yes
AVI	No	No	No	Yes	Yes

Since the EHRcom communication protocol is undefined yet, the statements given in Table IX are guesses based on the predecessor standard, i. e. ENV 13606-4. WADO and the DICOM network protocol can be used to communicate everything that can be encapsulated as a DICOM object, i. e. DICOM SR, DICOM images also PDF documents, for which an encapsulation into DICOM is defined in [DICOM Supplement 104 2005]. JPEG images and MPEG-2 videos can also be encapsulated in DICOM if the metadata required for the DICOM header is known, the latter [DICOM Supplement 42 2004] also being a recent addition to the DICOM standard.

In addition of the capabilities of the DICOM protocol, WADO supports the conversion of DICOM images to JPEG and DICOM SR documents to HTML and, optionally, to PDF and RTF. RID and XDS, being "content format agnostic", can in principle be used to communicate any type of record content as long as it can be identified with a MIME type. RID defines a number of default formats and requires any server to be able to convert documents to one of the default formats if requested by the client. Therefore, formats beyond CDA, PDF and JPEG (along with XHTML for the Retrieve Specific Information for Display service) are optional in RID and require either server side support for format conversion or client side support if documents are to be retrieved in the original format. Currently RID only specifies the use of CDA Release One (Level 1), but it is predicted that this will be updated to Release Two once this is finalised by the HL7 committee. Finally, XDS does not require, specify or limit any content format and can, therefore, be used for any content format. It is expected that additional constraints will be introduced

into XDS with the "document content profiles" in the future. EHRcom, CDA and DICOM SR are included in the list of formats that IHE mentions as formats that are expected to be used with XDS.

9.5 EHR Standard Harmonisation Efforts

It is clear from the above discussions that most of the EHR standards are currently evolving. There is also a clear trend towards a harmonisation and unification of formerly distinct EHR developments:

—The openEHR foundation has agreed on a Memorandum of Understanding with CEN/TC 251 to join the EHR-related efforts of both organisations. The openEHR archetype concept will be included in the revised version of the European Prestandard ENV 13606 which will be released as a European Standard.

—Furthermore, common reference model (classes and datatypes) based on the HL7 version 3 reference information model (RIM) will be developed, resulting in major changes in the openEHR architecture and EN 13606.

—Regarding standardisation, a national adoption of EN 13606-1 as the Australian EHR standard is envisioned and, finally, an adoption of EN 13606 by ISO/TC 215 as an International Standard is intended. At this point, openEHR would be a true superset of EN 13606 and possibly HL7 CDA.

—In addition to the harmonisation between EHRcom and openEHR, much effort has been invested into the harmonisation between EHRcom and HL7 version 3. The EHRcom reference model is based on the HL7 version 3 Reference Information Model (RIM) and can be expressed as an HL7 Domain Message Information Model (D-MIM), that is, as a specialisation of the HL7 RIM. The data types used in the EHRcom reference model are defined in CEN/TS 14796 [CEN prCEN/TS 14796 2003] and are a true subset of the HL7 version 3 data types. Furthermore, EHRcom attempts to implement the requirements for an EHR architecture described in ISO/TS 18308.

9.6 Analysis of Market Relevance of EHR Standards

Table XII summarises the market relevance of the EHR standards. It shows whether or not the standards are available as a final specification, whether implementations of the standard are known and whether or not implementations of the standard are available in off-the-shelf commercial products for the medical market. The final specifications are not yet available for EHRcom, CDA Release Two and XDS (which is currently a frozen draft for trial use and will be finalised in 2005). Prototype or project-based implementations are known to exist for all standards except EHRcom and XDS. IHE has made a public demonstration of the XDS integration profile in February, 2005. Commercial products are known to exist for CDA Release Two, for example Microsoft InfoPath 2003 [Infopath 2003] is based on a draft of the CDA Release Two standard, DICOM SR and RID.

Most but not all of the EHR standards are intended for an international market. MML, while probably applicable in many countries, has been developed specifically for the Japanese market. Finally, two of the EHR standards, WADO and DICOM SR focus on the medical imaging sector. While DICOM SR is a very generic content format that can be used to describe any kind of clinical data, the DICOM standard

Table XII. Market Relevance of EHR Standards

	CEN EHRcom	HL7 CDA R2	ISO WADO	DICOM SR	IHE RID	IHE XDS	MML
Final specification available	No	No	Yes	Yes	Yes	No	Yes
Implemented	No	Yes	Yes	Yes	Yes	Yes	Yes
Commercial products available	No	Yes	Yes	Yes	Yes	Yes	Yes
Intended for international market	Yes	Yes	Yes	Yes	Yes	Yes	No
Focussed on medical imaging sector	No	No	Yes	Yes	No	No	No

is traditionally closely coupled to the medical imaging sector, where DICOM is the prevalent format for an exchange and archival of medical images and related structured data. This is also reflected by the library of templates for DICOM SR which mostly contains templates for diagnostic imaging reports and measurements exchanged in the context of medical imaging and the post-processing of medical images.

10. CONCLUSIONS

The evaluation of the seven EHR standards reveals no clear "winner". The content format standards are surprisingly similar in concept and capabilities, using a two-level modelling approach with a simple reference model and constraint rules (archetypes, templates) for mapping clinical data onto the model. They differ in the progress achieved in the standardisation process: while DICOM SR already has a large library of standard templates encapsulating domain knowledge, EHRcom is still working on the archetype description language that will also be used for CDA. In principle, however, each of the content formats seems to be suitable for implementing electronic healthcare records.

The picture changes when one takes into account the market relevance of the standards. DICOM SR is clearly the most advanced EHR standard at the time being in terms of content structure, template library and access services. However, since the focus of DICOM is the medical imaging sector, SR is probably not the global solution for the EHR. While DICOM would probably be sufficiently powerful to be used as a comprehensive EHR standard, experience with practical implementation of SR over the last five years shows that the acceptance outside the medical imaging domain (i. e. Radiology, Cardiology etc.) is rather limited because vendors have no experience with the complex protocol and binary encoding rules used by the DICOM standard and the initial hurdle of getting a working understanding of the complex technical specification.

The Medical Markup Language seems to be a purely national initiative for the Japanese market. Since MML version 3 already uses the CDA format with a few extensions (local markup), it is likely that MML will be completely merged with the CDA specification once it supports templates.

The EHR access service standards vary significantly in approach, scope and the

progress achieved in the standardisation process. Again DICOM is the specification that has been stable for the longest time, with the network services being available since 1993, but it is not probable for DICOM to be successful in the EHR market. WADO only defines a supplementary protocol that needs to be combined with something else, in particular because WADO does not define any service for querying EHR content. The client needs to know in advance the unique identifiers of the document to be retrieved. The access services for EHRcom are currently under development, and for CDA it is not even clear if there will be standardised access services at all. The remaining standards are the two "content format agnostic" services defined by IHE: Retrieve Information for Display (RID) and Cross-enterprise Document Sharing (XDS).

While the intended use case for RID is inter-departmental (i. e. intra-enterprise) document exchange, it can easily be used in a cross-enterprise setting using the same approach that is also used in XDS, i. e. using the actors of the IHE PIX integration profile.

The XDS integration profile defines all access services needed for an EHR and has been designed for cross-enterprise use. The recent XDS extensions such as Cross-Enterprise User Authentication [IHE-XUA 2005] and the first set of document content profiles [IHE-DDS 2005; 2005] address the most important issues that need to be covered in order to make XDS a viable option for EHR implementation projects. There is still a number of open issues such as access control and document content profiles for non-imaging domains. However, there seems to be significant interest in XDS, both from the commercial side and from national EHR projects such as the "Canada Health Infoway" [Canada Healthh Infoway], which means that a successful market uptake seems likely. The first commercial products implementing central IHE XDS actors are available already.

If we think of a long-term solution, e. g., for the next five years, other "candidates" seem possible, although the development is still too much in flux today to forecast any details:

—*XDS with structured content in CDA or EHRcom format and image access through WADO (XDS-I):* This combination of standards, together with some XDS document content profiles that could specify CDA templates for specific use cases, would provide a comprehensive cross-enterprise EHR solution.

—*EHRcom as a comprehensive EHR solution:* Since EHRcom is attempting to define a comprehensive EHR architecture, the complete set of EN 13606 parts 1-5 will also provide a comprehensive cross-enterprise EHR architecture, based on HL7 messages. It is very difficult to forecast the market uptake of EHRcom. In the past, the success of most standards developed by CEN/TC 251 was quite limited, except for the EDIFACT messages that are used in some of the Scandinavian countries and in the UK. The convergence of EHRcom, openEHR and HL7 makes a success of the new standard certainly more likely compared to older CEN works such as the pre-standard ENV 13606.

—*CDA as a comprehensive EHR solution:* The scope of HL7 encompasses all data exchanges "that support clinical patient care and the management, delivery and evaluation of healthcare services." Although it is not clear whether or not HL7 will develop a comprehensive EHR architecture based on CDA, it is likely that

CDA-related HL7 version 3 messages will be defined in the future. It remains to be seen whether this provides an alternative to the architectures described above.

One final issue to be addressed is the following: given the large number of EHR standards, conformance to one of these standards or implementing a combination of them will not solve the interoperability problem; there will always be some healthcare institutes using a different, incompatible EHR standard. It is not realistic to expect all the healthcare institutes to reach a consensus on a single universal standard. Therefore in the longer run, true interoperability of EHRs will only be possible by providing semantic interoperability.

Semantic interoperability is the ability for information shared by systems to be understood at the level of formally defined domain concepts so that the information is computer processable by the receiving system [ISOTS18308]. In other words, semantic interoperability requires the semantics of data to be defined through formally defined domain specific concepts. An important initiative in this respect is the Artemis project [ARTEMIS]. The Artemis middleware enables healthcare institutes to exchange Electronic Healthcare Records in an interoperable manner through semantically enriched Web services and semantic mediation. Further details of this approach, which is currently at the research and development stage, are discussed in [Dogac et al. 2005; Dogac et al. 2004; Bicer et al. 2005a; 2005b; Bicer et al. 2005c; Aden et al. 2004; Aden and Eichelberg 2005; Boniface and Wilken 2005].

REFERENCES

Aden, T. and Eichelberg, M. 2005. Cross-enterprise search and access to clinical information based on IHE Retrieve Information for Display. In *Proceedings of EuroPACS-MIR 2005.*

Aden, T., Eichelberg, M., and Thoben, W. 2004. A fault-tolerant cryptographic protocol for patient record requests. In *Proceedings of EuroPACS-MIR 2004 in the enlarged Europe.* Trieste(Italy).

ADL 1.2. Archetype Definition Language (ADL) 1.2 draft. http://www.openehr.org/drafts/ADL-1_2_draftF.pdf.

ANSI. American National Standards Institute. http://www.ansi.org/.

Araki, K., Ohashi, K., Yamazaki, S., Hirose, Y., Yamashita, Y., Yamamoto, R., Minagawa, K., Sakamoto, N., and Yoshihara, H. 2000. Medical Markup Language (MML) for XML-based Hospital Information Interchange. *Journal of Medical Systems 24,* 3, 195–211.

ARTEMIS. ARTEMIS Project. http://www.srdc.metu.edu.tr/webpage/projects/artemis.

Beale, T. 2002. Archetypes: Constraint-based Domain Models for Future-proof Information Systems. In *OOPSLA 2002 workshop on behavioural semantics.* http://www.deepthought.-com.au/it/archetypes/archetypes_new.pdf.

Beale, T. and Heard, S. Archetype Definitions and Principles. http://www.openehr.org/-repositories/spec-dev/latest/publishing/architecture/archetypes/principles/REV_HIST.html.

Bicer, V., Laleci, G., Dogac, A., and Kabak, Y. 2005a. Artemis Message Exchange Framework: Semantic Interoperability of Exchanged Messages in the Healthcare Domain. *ACM Sigmod Record 34,* 3 (September).

Bicer, V., Laleci, G., Dogac, A., and Kabak, Y. 2005b. Providing Semantic Interoperability in the Healthcare Domain through Ontology Mapping. In *Proceedings of eChallenges 2005.* Ljubljana, Slovenia.

Bicer, V., Kilic, O., Dogac, A., and Laleci, G. 2005c. Archetype-based Semantic Interoperability of Web Service Messages in the Healthcare Domain. *International Journal on Semantic Web and Information Systems.* Submitted for Publication.

Boniface, M. and Wilken, P. 2005. ARTEMIS: Towards a Secure Interoperability Infrastructure for Healthcare Information Systems. In *Proceedings of Healthgrid2005*. Oxford, UK.

Canada Healthh Infoway. Canada Health Infoway. http://www.canadahealthinfoway.com/home.-php?lang=en.

CBC Archetype. Complete Blood Count Archetype ADL Definition. http://www.openehr.-org/repositories/archetype-dev/adl_1.1/adl/archetypes/openehr/ehr/entry/openehr-ehr--observation.haematology-cbc.draft.adl.html.

CEN ENV 12265 1997. *ENV 12265:1997, Medical Informatics - Electronic Healthcare Record Architecture, CEN/TC 251 Health Informatics, European Committee for Standardization*. Brussels, Belgium.

CEN ENV 13606 2000. *CEN ENV 13606:2000, Medical Informatics - Electronic healthcare record communication, European Prestandard, CEN/TC 251 Health Informatics, European Committee for Standardization*. Brussels, Belgium.

CEN prCEN/TS 14796 2003. *prCEN/TS 14796, Health informatics - Data types, CEN/TC 251 Health Informatics, European Committee for Standardization*. Brussels, Belgium.

CEN prEN 14822-1 2003. *CEN prEN 14822-1 Health informatics - General purpose information components, CEN/TC 251 Health Informatics, European Committee for Standardization*. Brussels, Belgium.

CEN TC251. CEN TC/251 European Standardization of Health Informatics. http://www.centc251.org/.

Channin, D. S., Siegel, E. L., Carr, C., Sensmeier, J., Henderson, M., Behlen, F. M., Parisot, C., and Siegel, E. L. 2001b. Integrating the Healthcare Enterprise: a primer. Part 4–5. *Radiographics 21,* 6 (Nov–Dec), 1597–1608.

Channin, D. S., Siegel, E. L., Parisot, C., Wanchoo, V., and Leontiev, A. 2001a. Integrating the Healthcare Enterprise: a primer. Part 1–3. *Radiographics 21,* 5 (Sep–Oct), 1339–58.

CLUNIE, D. A. 2000. *DICOM Structured Reporting*, 1 ed. PixelMed Publishing, Bangor PA, USA.

DICOM 2004. *NEMA Standards Publication PS 3.x, Digital Imaging and Communications in Medicine (DICOM), National Electrical Manufacturers Association*. Rosslyn VA, USA.

DICOM Supplement 104 2005. DICOM Encapsulation of PDF Documents, DICOM Standards Committee, Working Group 6, Final Text. ftp://medical.nema.org/medical/dicom/final/sup104_ft.pdf.

DICOM Supplement 42 2004. MPEG2 Transfer Syntax, DICOM Standards Committee, Working Group 13 Visible Light, Final Text. ftp://medical.nema.org/medical/dicom/final/sup42_ft.pdf.

DICOM Supplement 85 2004. Web Access to DICOM Persistent Objects (WADO), Joint DICOM Standards Committee / ISO TC215 Ad Hoc WG on WADO, Final Text. ftp://medical.nema.org/medical/dicom/final/sup85_ft.pdf.

DICOM Supplement 99 2005. Extended Negotiation of User Identity, DICOM Standards Committee, Working Group 14, Final Text. ftp://medical.nema.org/medical/dicom/final/sup99_ft.pdf.

Dierks, T. and Allen, C. 1999. The TLS Protocol Version 1.0, IETF RFC 2246. http://www.ietf.org/rfc/rfc2246.txt.

Dogac, A., Laleci, G., Kirbas, S., Kabak, Y., Sinir, S., Yildiz, A., and Gurcan, Y. 2005. Artemis: Deploying Semantically Enriched Web Services in the Healthcare Domain. *Information Systems Journal (Elsevier) special issue on Semantic Web and Web Services*. To appear.

Dogac, A., Laleci, G. B., Kabak, Y., Unal, S., Beale, T., Heard, S., Elkin, P., Najmi, F., Mattocks, C., and Webber, D. 2004. Exploiting ebXML Registry Semantic Constructs for Handling Archetype Metadata in Healthcare Informatics. *International Journal of Metadata, Semantics and Ontologies*. To appear.

ebRIM. ebXML Registry Information Model v2.5. http://www.oasis-open.org/committees/-regrep/documents/2.5/specs/ebRIM.pdf.

ebRS. ebXML Registry Services Specification v2.5. http://www.oasis-open.org/committees/-regrep/documents/2.5/specs/ebRIM.pdf.

EHRcom 2004. *rEN 13606-1, Health informatics - Electronic health record communication - Part 1: Reference model, Draft for CEN Enquiry, CEN/TC 251 Health Informatics, European Committee for Standardization.* Brussels, Belgium.

Eichelberg, M., Poiseau, E., B.Wein, B., and Riesmeier, J. 2003. IHE Europe: extending the IHE initiative to the European healthcare market. In *Medical Imaging 2003: PACS and Integrated Medical Information Systems: Design and Evaluation*, H. K. Huang and O. M. Ratib, Eds. Vol. 5033. SPIE, 1 – 10.

FREEebXML. freebXML (OMAR) implementation. http://ebxmlrr.sourceforge.net/3.0/index.-html.

Guo, J., Takada, A., Tanaka, K., Sato, J., Suzuki, M., Suzuki, T., Nakashima, Y., Araki, K., and Yoshihara, H. 2004. The Development of MML (Medical Markup Language) Version 3.0 as a Medical Document Exchange Format for HL7 Messages. *Journal of Medical Systems 28,* 6, 523–533.

HIMSS. Healthcare Information and Management Systems Society. http://www.himss.org/-ASP/index.asp.

HL7. Health Level 7 (HL7). http://www.hl7.org.

HL7 CDA Implemantations. HL7 Clinical Document Architecture Implementations. http://-www.hl7.org.au/CDA.htm.

HL7CCOW. HL7 CCOW (Clinical Context Object Workgroup) Standard. http://www.hl7.org.-au/CCOW.htm.

HL7CDA Release 1.0 2000. CDA, HL7 Clinical Document Architecture - Release 1.0, ANSI Standard.

HL7CDA Release 2.0 2004. CDA, HL7 Clinical Document Architecture - Release 2.0. http:/xml.-coverpages.org/CDA-20040830v3.pdf.

HL7RIM. HL7 Reference Information Model. http://www.hl7.org/library/data-model/RIM/modelpage_non.htm.

HL7v2.5 2000. *HL7, Application Protocol for Electronic Data Exchange in Healthcare Environments, Version 2.5, ANSI Standard.* Ann Arbor MI, USA.

HL7V3. HL7 Version 3 Message Development Framework. http://www.hl7.org/library/mdf99/-mdf99.pdf.

Hussein, R., Engelmann, U., Schröter, A., and Meinzer, H.-P. 2004a. DICOM structured reporting: Part 1. Overview and characteristics. *Radiographics 24,* 3 (May–Jun), 891–96.

Hussein, R., Engelmann, U., Schröter, A., and Meinzer, H.-P. 2004b. DICOM structured reporting: Part 2. Problems and challenges in implementation for PACS workstations. *Radiographics 24,* 3 (May–Jun), 897–909.

IHE-DDS 2005. IHE IT Infrastructure Technical Framework, Supplement 2005-2006, Document Digital Signature, Public Comment Version, IHE IT-I Technical Committee. http://www.ihe.net/Technical_Framework/upload/IHE_ITI_TF-Supplement_Digital_Signature-PC_2005-06-15.pdf.

IHE IT-IP. IHE IT Infrastructure Integration Profiles. http://www.rsna.org/IHE/tf/ihe_tf_index.-shtml.

IHE-XDS 2004. *IHE IT Infrastructure Technical Framework, Cross-Enterprise Clinical Documents Sharing (XDS), IHE ITI Technical Committee.* Version 3.0.

IHE-XDS-I 2005. IHE Radiology Technical Framework, Supplement 2005-2006, Cross-enterprise Document Sharing for Imaging (XDS-I), Public Comment Version, IHE Radiology Technical Committee. http://www.ihe.net/Technical_Framework/upload/IHE_RAD-TF_Suppl_XDSI_PC_2005-06-15.pdf.

IHE-XUA 2005. IHE IT Infrastructure Technical Framework, Supplement 2005-2006, Cross-Enterprise User Authentication (XUA), Public Comment Version, IHE IT-I Technical Committee. http://www.ihe.net/Technical_Framework/upload/IHE_ITI_TF-Supplement_Cross-Ent_User_Authentication_PC_2005-06-15.pdf.

IHEMS. IHE Mission Statement. http://www.rsna.org/IHE/mission.shtml.

IHETF. IHE Technical Framework. http://www.rsna.org/IHE/tf/ihe_tf_index.shtml.

Infopath 2003. Microsoft InfoPath 2003. http://www.microsoft.com/office/infopath/prodinfo/-default.mspx.

ISO/DIS 17432 2004. *ISO/DIS 17432, Health informatics - Messages and communication - Web access to DICOM persistent objects, Draft International Standard, International Organization for Standardization (ISO)*. Geneva, Switzerland.

ISOTR18307 2001. *ISO/TR 18307, Health informatics - Interoperability and compatibility in messaging and communication standards - Key characteristics, Technical Report, International Organization for Standardization (ISO)*. Geneva, Switzerland.

ISOTR20514 2004. *ISO/CD TR 20514, EHR, definition, scope and context, Committee Draft, Technical Specification, International Organization for Standardization (ISO)*. Geneva, Switzerland.

ISOTS18308. ISO TC/215, International Organization for Standardization, Health Informatics, ISO TS 18308:2003. secure.cihi.ca/cihiweb/en/downloads/infostand_ihisd_isowg1_mtg_denoct-_contextdraft.pdf.

Kalra, D. 2004. ICEN prEN 13606, draft standard for Electronic Health Record Communication and its introduction to ISO TC/215, CEN/TC 251 WG I Document N04-52.

Kerberos. Kerberos: The Network Authentication Protocol. http://web.mit.edu/kerberos/www/.

LOINC. Logical Observation Identifiers Names and Codes (LOINC). http://www.loinc.org/.

MML 2003. Medical Markup Language (MML) Specification Version 3.0, MedXML Consortium. http://www.medxml.net/E_mml30/.

NAICS. North American Industrial Classification Scheme (NAICS) codes. http://www.naics.com.

NEMA. National Electrical Manufacturers Association (NEMA). http://www.nema.org/.

OpenEHR. openEHR Community. http://www.openehr.org/.

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

SAML Overview 2005. Security Assertion Markup Language (SAML) 2.0 Technical Overview, Working Draft 06. http://www.oasis-open.org/committees/download.php/12938/sstc-saml-tech-overview-2.0-draft-06.pdf.

SAML Profiles 2005. Profiles for the OASIS Security Assertion Markup Language (SAML) V2.0, OASIS Standard. http://docs.oasis-open.org/security/saml/v2.0/saml-profiles-2.0-os.pdf.

SNOMED. SNOMED (The Systematized Nomenclature of Medicine) Clinical Terms. http://www.snomed.org/snomedct_txt.html.

UNSPSC. Universal Standard Products and Services Classification (UNSPSC). http://eccma.org/unspsc.

WSDL. WSDL: Web Service Description Language. http://www.w3.org/TR/wsdl.

Received June 2005