A flexible model for the delivery of multi-facet information in patient–centric Healthcare Information Systems

Iraklis Varlamis
Athens University of Economics and Business
Department of Computer Science
varlamis@aueb.gr

Abstract
This paper suggests a new approach for the development of healthcare information standards, which is based on widely used and open frameworks. The paper attempts a review of existing standards for healthcare information, analyses their deficiencies and focuses on the need for interoperability. Healthcare information, in order to be useful, has to be well formed, valid and flexible. Healthcare information standards are the pre-requisites for well-formedness and validity of information. Flexibility expects “plug-and-play” information, which will be operable in any system, any time and any place. A standard that will be accepted world-wide and will cover all possible aspects of healthcare information needs sounds both infeasible and inconvenient due to its size and complexity. A more practicable solution is to provide an interconnection mechanism on top of all existing and future standards. This work specifies the two fundamental problems of information interoperability, which are structure and semantics, and suggests a mechanism that facilitates the integration of existing information, the mutation and transfer of information between healthcare information systems and eases interoperability.

Keywords: Syntactic and semantic interoperability, Healthcare information systems, Multi-facet information

1. Introduction
The plurality of healthcare solutions, the advent of internet and the abundance of health information, the increased competition among healthcare companies, and the increased mobility of patients has significantly affected the healthcare industry. As stated by Moser [Moser, 1992], patients are not solely dependent on their doctor anymore, they are aware of their healthcare record, they participate on decisions and choose among available solutions, and they slowly transform into healthcare consumers. In this competitive market companies create flexible eHealth networks, which offer their services and support patients world-wide. eHealth coalitions imply integration of information systems and interoperability of exchanged medical information. In the same time, healthcare information should be available in multiple formats and granularities in order to be useful for professionals, patients and healthcare information systems. Additional parameters such as access rights, digitized data and encoding information etc must be easily attached and detached from the core information.

The majority of existing health information systems has been designed to serve only one or few departments within a healthcare institute. Health information produced in a clinic for a patient (i.e. examination results, diagnoses etc) cannot be integrated to the information systems of other clinics and becomes useless. Patients that change healthcare providers (e.g. when they move to a new place or choose a different insurance company) are obliged to undergo the same medical examinations multiple times. Their healthcare record is distributed into different healthcare institutes, parts of the record are overlapping and even contradicting and it is very difficult for clinicians to maintain a complete clinical history of a patient [Apostolakis 2002].

The fundamental concept behind this work is that all participants of the healthcare community (hospitals, clinics, healthcare product providers, insurance companies, patients etc) should be able to work either individually or jointly, accessing a common patient healthcare record and to communicate with messages that share the same semantics, in order to provide each patient with the best available care anywhere in the world. This states the need for interoperability of healthcare information: either it is the healthcare record of patients or the communication messages between information systems and healthcare devices. This is the best solution for small scale companies or organizations that cannot cover a country-wide or continent-wide area, and for patients that will be able to join any healthcare network, choose among the available solutions and make use of the services offered. They can also be

eJETA.org Special Issue. Interoperability & Security in Medical Information Systems
informed on their healthcare record which will be available in various detail levels. The key for this solution is to achieve interoperability and quality of healthcare information and services [Apostolakis & Valsamos, 2005].

In the same concept, interoperable information should be machine readable and understandable by all cooperating information systems. For this reason, we should devise a model to describe healthcare information that will be rich in semantics and sturdy in structure. We should also set up a mechanism that will interpret the semantics and re-format the structure of any piece of information transferred from one system to another. Any future information system can be easily incorporated into the existing structure by building its own interpretation mechanism.

The following section presents an overview of existing standards and nomenclatures used in healthcare. Section 3, discusses some mentionable works on interoperability of healthcare information. The section presents both semantic and structural aspects of interoperability. Section 4 discusses the proposed solution and the merits that arise from the use of open standards and architectures. Section 5 provides an example case of accessing interoperable healthcare information.

2. Healthcare standards

2.1 Healthcare messaging

For exchanged healthcare information (messages), Health Level 7 (HL7) Messaging Standard is the most widely accepted messaging standard for communicating clinical data and it is supported by every major medical informatics system vendor in the US. Unfortunately, the standard has no explicit information model, has vague definitions for many data fields, it contains many optional fields and many fields with undefined cardinality of values [Neotool]. It offers great flexibility, but requires strict agreements among healthcare systems to achieve interoperability.

For the healthcare information that is stored for every patient the problem is twofold: what to store and where to store it. The Electronic HealthCare Record (EHR) comprises data in the computer systems of all healthcare organizations or providers who care for a patient (hospitals, physical therapists, pharmacists, or consulting physicians). A number of standardization efforts ([EHRcom], [openEHR] and [HL7 V3]) address the problem of fitting the Electronic Record of a patient, which is stored in the individual information system of a doctor or clinic to a HealthCare Record that can be used by anyone. For the same reason, the Medical Records Institute [MRI] distinguishes five levels of an Electronic HealthCare Record: a) paper based record, b) computerized medical record, c) interoperable electronic medical record, d) patient-centered record and finally e) general health information relevant to the patient’s record.

The Medical Data Interchange Standard-MEDIX [Harrington, 1991] has been developed by IEEE to support information exchange between healthcare information systems. It defines the structure of exchanged messages and technical issues such as emailing, resource identification, file transfer etc. EDI is widely used for document exchange between applications upon agreement on the message format. United Nation’s [EDIFACT] and ANSI [ASC X12] have hierarchical structure and allow the composition of complex structure from simpler ones. The standard includes a header for identifying the message sender. DICOM [NEMA] standard was initially designed for the recording and exchange of radiology images. It is used by filing systems such as PACS (Picture Archiving and Communications Systems) and focuses on computer-medical device interfacing.

In order to cover patient mobility, standards have been developed that identify patients, healthcare providers, healthcare places and products. A complete healthcare record should contain such information, in order to identify the owner of the record and to have complete coverage on the patient’s history.

Patient identification standards facilitate providers to maintain a single EMR for each patient and retrieve it from any information system. Identification standards allow patients to access their own record and protect it from unauthorized access. Nowadays, the Universal Healthcare Identifier-UHID-E1734 [ASTM, 1995] developed by ASTM (American Society for Testing and Materials) is a widely accepted solution. Other suggestions include the use of Social Security Number, biometric patient data etc. Provider identification standards, such as National Provider Identifier (NPI), define unique codes for every healthcare provider, and encapsulate all the remaining information. The Health Industry
Number (HIN) uniquely defines hospitals, pharmacies, private doctors and clinics, providers etc and supports the referencing of departments inside a healthcare institute. Finally, the Labeler Identification Code (LIC), EAN/UCC and NCDPD standards identify medical products and providers.

2.2 Healthcare terminologies

a) The International Classification of Diseases (ICD) [Gersenovic, 1995] of World Health Organization gives a unique code to every disease. The classification has been accepted world-wide by healthcare and insurance organizations, researchers and practitioners. The ICD codes appear in a wide range of medical documents, treatment records, patients’ records etc. Every code describes the affected part of the body, the reason for the disease etc.

b) The International Classification in Primary Care - ICPC-2 [ICPC2] encodes the interaction between the patient and the doctor in Primary Healthcare Units during a medical incident or in a series of medical events (i.e. visits at home). In an incident, both the problem and the symptoms are described and the reason for medical treatment, the diagnosis and the medical care are recorded. The standard has been mapped to the ICD structure in an effort to standards integration.

c) The Read Codes (or Clinical Terms) [Read Codes] describe the medical treatment of patients using a hierarchy of terms [O’Neil et al., 1995], which comprise symptoms, exams, diagnosis, treatments and medicines. The Centre for Coding and Classification of British National Health Systems uses Health Codes to describe cases ranging from a medical incident to an Electronic Healthcare Record.

d) The SNOMED (Systematized Nomenclature of Medicine) [Cote et al, 1993] standard defines a hierarchical terminology that covers all aspects of HER. The standard supports cross references to more than one code per concept. The SNOMED Clinical Terms unifies SNOMED and Read Codes and comprises 300,000 concepts and 1,000,000 relations among them. It also matches ICD-10 codes and most of the LOINC terms.

e) The Diagnosis Related Group (DRG) standard [Feinstein, 1987] has been created by Health Care Finance Administration in order to correlate diseases and treatment costs and serve hospitals and insurance companies. It classifies ICD diseases in 23 main categories and many more subcategories in order to facilitate and standardize treatment cost calculation.

f) The Anatomic Therapeutic Chemical [ATC] is a five level hierarchical classification of drugs used by many European countries.

g) The LOINC standard [McDonald, 1995] has been implemented, in order to cover medical examination standardization needs. Was designed to support HL7 and has been adopted by DICOM. Comprises more than 34,000 exam codes that define the kind of exam, the type of measurement, the duration of the exam, the scale etc.

h) Finally, Diagnostic and Statistical Manual for Mental Disorders (DSM) standard provides a classification of mental disorders into 5 axis and is compliant to the ICD standard.

Figure 1 presents the various standards, their relevance and the mappings that have been defined between some of them.

![Figure 1. Healthcare semantics and terminology standards](image)

As it is depicted in the figure, and it is made obvious in the above analysis, there exists a global effort to integrate all the major healthcare standards. However, integration is perceived as appropriate linkage
between existing standards. Healthcare organizations prefer to provide a mapping between their and others’ standards than to agree on a single standard that will cover everything.

Table 1 gives a summary of the most important messaging and terminology standards, their primary aim and the underlying technologies.

<table>
<thead>
<tr>
<th>Messaging Standards</th>
<th>Aim</th>
<th>Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>HL7 CDA, EHRcom</td>
<td>Clinical data interchange</td>
<td>XML</td>
</tr>
<tr>
<td>ASTM CCR, CONTSYS</td>
<td>Continuity of Care Record</td>
<td>XML</td>
</tr>
<tr>
<td>ANSI X12</td>
<td>Financial data, Billing</td>
<td>EDI</td>
</tr>
<tr>
<td>DICOM</td>
<td>Images</td>
<td>DICOM</td>
</tr>
<tr>
<td>openEHR</td>
<td>Define demographics, clinical workflow etc. Assist searching, human communication, improve data validity, increase reusability</td>
<td>ADL</td>
</tr>
<tr>
<td>IEEE</td>
<td>Data interchange</td>
<td>MEDIX</td>
</tr>
<tr>
<td>ISO TC215 standards</td>
<td>Interoperability of telehealth and telelearning systems</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Terminology Standards</th>
<th>Aim</th>
<th>Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOINC</td>
<td>Laboratory observations</td>
<td></td>
</tr>
<tr>
<td>ICD, I CPC-2, ReadCodes, SNOMED</td>
<td>Clinical information</td>
<td>OWL, RDF or extensions</td>
</tr>
<tr>
<td>ATC, NCPDP</td>
<td>Drugs</td>
<td></td>
</tr>
<tr>
<td>ICD-9CM, DRG</td>
<td>Billing, Diagnosis</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Healthcare messaging and terminology standards and technologies

3. Research work on information interoperability

According to the ideal scenario, every healthcare participant (institutes, organizations, companies and individuals) shares the same information infrastructure, which follows widely accepted standards. Integrated healthcare information systems are intended to provide each patient with a globally accessible medical record. This medical record contains information that varies from identification data to sexual diseases and serious illnesses. The security of transferred medical data is another high priority issue to be considered by the designers of medical information systems.

In reality, despite the effort of healthcare committees to provide standards [NHII] for healthcare information interoperability, an exchange of well-structured and machine-processable electronic healthcare records has not been achieved yet in practice [Dogac et al, 2006]. Interoperable healthcare information systems must be built upon a sturdy information structure using cohesive semantics as glue [WHO] between different approaches and mapping mechanisms as information bridges [Nainil & Chheda, 2007].

![Interoperability of healthcare information](image)
3.1 The need for semantics

Despite the abundance of standards, protocols and structural models, health care information currently lacks of semantics. In a realistic healthcare setting today, the exchanged message instances are EDI or XML, and not messages conforming to an ontology. As a result, information is machine-readable but not understandable and human intervention is required in every transaction. In order to automatically understand documents we need terminology and semantics. A lot of terminology standards have been developed to express in a systemic way diagnostic and clinical processes, to describe healthcare information in a "commonly accepted language" and to avoid vague definitions and errors. More specifically:

Healthcare information systems are able to store or exchange well formed and valid patient records and messages that comply with standards with rich semantics. Healthcare documents complying EDI, DICOM, ISO, HL7 etc. can be automatically converted to any other standard if a mapping mechanism is defined. In order to increase the efficiency of mapping and improve information interoperability we should enrich structural matching with semantics [Bicer et al, 2005].

The HL7 Clinical Document Architecture [CDA] is an XML-based mark-up standard intended to specify the encoding, structure and semantics of clinical documents for exchange. It is based on the HL7 Reference Information Model (RIM) and the HL7 Version 3 Data Types, though can be used independently of any HL7 Version 3 messaging. The architecture supports narrative text in order to ensure that the content will be human-readable, contains structure, and most importantly, allows for the use of codes (such as from SNOMED and LOINC) to represent concepts.

The Continuity of Care Record [CCR] is a standard developed jointly by the American Society for Testing and Materials (ASTM) and other health informatics vendors and associations in the United States. It is designed to facilitate creation of medical documents that contain the most relevant and timely core health information about a patient, and to assist electronic transfer from one care provider to another. It contains patient demographics, insurance information, diagnosis and problem list, medications, allergies, care plan and other critical information that can be lifesaving if available at the time of clinical encounter. Documents that follow the CCR standard can be easily created by a physician using an electronic health record (EHR) system.

Aiming in reusability of components, openEHR introduces the concept of Archetypes and Templates [Heard & Beale, 2005] that can be used in the description of patient demographics, clinical workflow etc. The reusable components are described using Archetype Definition Language (ADL), which uses XML and UML notation to define concepts of a healthcare information system as with any other information system.

4. The proposed solution

The review of available standards shows that healthcare information systems are able to exchange data only upon agreement on the standards and semantics and stated the need for exchange of meaningful clinical information among healthcare institutes. This summary presented the necessities and deficiencies of healthcare information systems and highlighted the roadmap towards efficient eHealth networks.

The first step is to conclude in a minimum set of standards that cover all healthcare activities. The standards should be open to extensions or modifications, must have a concrete structure and be rich in semantics. The mapping mechanisms that will be developed on top of the information will carry domain knowledge and will be able to efficiently match semantics from two or more standards, to combine structured information from many sources and produce healthcare information for individuals and organizations.

The use of hierarchical information structures and XML related technologies seems to be the preferred solution for healthcare information standards. Information can be semantically annotated using one or more relevant OWL ontologies, which provide the nomenclature and conceptual model for interpreting and reasoning with the concept. The ICD-10 standard for clinical documents has already been the backbone for the integration of clinical data. Most standards in this category are aligned to ICD-10 thus facilitating the semantic interoperability of information. OWL or RDF technologies can be the
backbone for describing healthcare semantic knowledge and other semantic technologies, such as [SWRL] and [RDQL], can be used to provide ontology alignment and information matching.

The final step is to build the mechanisms that will take this well structured (XML) information and its semantics (RDF) as input and will produce new information as output. The Semantic Web and Web Services are the ideal solution for this conversion [Apostolakis & Valsamos, 2006]. Information exchange between cooperating systems can be achieved by sending information through the appropriate web service, defining the source and target syntactic and semantic standards and receiving the processed information on target. Definitely, information exchange and access must be performed using validation and authentication mechanisms, which should be supported by the web service provider and the cooperating information systems but this is outside of the scope of this paper.

Figure 3 that follows, presents the proposed architecture with the information sources, formats and transformations being annotated with the respective technology acronyms, which will be presented in the following paragraphs.

![Figure 3. Architecture for the interoperability of healthcare information](image)

Different content delivery services may access the same healthcare records and produce a filtered output. In order to ensure privacy protection, the access to the web services must be restricted only to authorized personnel (i.e. doctors, nurses). Every user, depending on his/her level and position, will be eligible to access specific services and receive filtered information. This means that the original healthcare record information will remain intact in a central repository, and different filtered versions will be delivered upon request to the accredited users via service calls. Another merit of this approach is that there will be no loss of information, rather hiding or mutation of the information being delivered.

### 4.1. Syntactic interoperability

The syntactic interoperability is crucial for information integration. As it is depicted in table 1, the majority of messaging standards use XML or EDI technologies to represent data. Even for those that have their proprietary structure, such as ADL or DICOM, the transformation to XML has already been specified. An extended research on available mapping mechanisms gives evidence on the following mappings:

- HL7’s EDI to XML: The open-source programming library from HL7, namely HL7 application programming interface [HAPI] can be used in transforming the EDI messages into their XML representations
- ADL to XML: serialization to XML and other ADL transformation are available thanks to The openEHR Eiffel Reference Implementation Project [openEHR Impl].
- **DICOM to XML**: DICOM’s Structured Reporting (SR) standard has been described in DTD and XML-Schema notation by researchers [Xiaozhen & Zhihong 2005] and several APIs have been developed, such as [openDICOM.NET], which supports DICOM as XML. Since the programming interfaces are available, and the logic of mapping between standards is clear, we can easily build web services responsible for transforming exchanged messages. Each Syntactic Conversion Web Service (SCWS) should be aware of the source and target electronic document structure (i.e. of the two XML-Schema files) and of the mappings between information elements. For example, a message produced by the Medical Lab IS using HL7’s EDI format will be forwarded to the appropriate semantic web service. There EDI will be transformed in XML and then forwarded to another service that will transform HL7’s XML to CCR XML. Then the document will be sent to its destination, which is the Hospital IS and will be incorporated to patients’ medical record. An encryption and decryption procedure will take place right before sending the source information and right after receiving the processed document.

**Figure 4. Web Services for syntactic interoperability**

### 4.2. Semantic interoperability

The overview has shown a wide variety of standards and terminologies used in every healthcare activity. Although these terminologies do not align with each other, several research and standardization efforts have been done to this direction. In order to achieve fully interoperable information we need another set of Web Services that will handle semantic alignment (see Semantic Matching Web Services – SMWB – in Figure 3). Most of the existing terminologies and healthcare standards have their Semantic counterpart.

- **LOINC**: The Regenstrief LOINC Mapping Assistant [RELMA] is an effort to facilitate searches through the LOINC database and to assist efforts to map local codes to LOINC codes. A mapping from LOINC to the CPT standard used for billing is also under development.

- **SNOMED**: Other European Projects such as [RIDE] support the incorporation of references to SNOMED codes into the LOINC or HL7 information.

- **HL7**: Artemis project [Bicer et al, 2005b] offers a mechanism for mapping HL7 v2 messages to HL7 v3 using semantics expressed in OWL syntax.

In general, RDFS and OWL are two languages that can be used to represent any of the existing terminology standards. They offer the mechanisms for defining the matching rules between terms in two different terminologies. For the semantic interoperability of information, we need services that are aware of the terminology and structure of documents. Semantic web services will be able to pump information from the source document, translate it to the new terminology and fit it into the target structure. For example, in order to incorporate laboratory results into the EHR, first the proprietary local codes should be matched to the 6 LOINC attributes [Lau et al, 2000], and consequently the LOINC-ified information can be incorporated into an HL7 message. The HL7 message will be transformed into a CCR document that will be imported into patient’s EHR. The procedure is depicted in Figure 5.
4.3. Multi-facet healthcare information

A major advantage from the use of web services and open information standards is the ability to reproduce information in various formats and levels of detail. Although the EHR of a patient is capable in covering all aspects of the healthcare process, only small portions of it are required in each aspect, and these portions must be transformed to an appropriate format that can be read by humans or utilized by computer systems. The information that is important for a practitioner in the case of an emergency differs from that needed to an insurance company and both are complicated and useless for the patient, who only wants a medical advice or a drug prescription. A doctor needs information in a browsable format (i.e. in XML) so that she can have an overview of the patient’s health state and in the same time be able to drill down into the details of a patient’s record, whereas a patient will be happy with a brief and clean diagnosis and medical prescription. Similarly, we should intentionally hide information from an unauthorized application or person by removing part of the original EHR data before presentation.

The suggested technologies (XML, RDF, OWL) allow the creation of views over the EHR information and thus are appropriate for producing multi-facet information. Once again, transformation web services (both syntactic and semantic ones) will be responsible for the final output (see Figure 3).

4.4. The merits of information interoperability

First of all, a system that builds upon XML and Semantic Web Services has all the merits of those two technologies [Pruitt et al, 2000], [Goble et al, 2001] [McIlraith et al 2001]: ease of representation, analysis tools, traceability, creation of views, collaboration, information and knowledge repositories. Especially for healthcare information the gains are multiple and are summarized in the following:

**Information is accessible everywhere.** Patients, doctors, hospitals are able either to recover the patient’s EHR from everywhere and import it into their information system, or to send the medical examination results and diagnosis as a message to the identification server and update the patient’s HER automatically.

**The EHR details are encapsulated and available to authorized users.** Since the communication with the EHR is performed via the identification server, only authorized users are able to access a patient’s record and modify or update its contents.

**Personal information is secure.** Information is encrypted before being transferred to the server thus guaranteeing protection of sensitive personal data.

**Healthcare information is easy to read and interpret.** The original results of a medical examination are valuable to a doctor, but useless to the common patient. With the use of semantics, the value of “140/90 mmHg” for blood pressure can be converted to “High” and marked with red color into the PDF document that is sent to the patient, thus increasing readability of HER contents. Only the information which is delivered to the user is transformed without affecting the original healthcare record information, so there is no information loss.

**Advanced knowledge management is supported.** Information with semantics can be used as a knowledge base for decision support, for population statistics or healthcare planning tasks [Varlamis & Apostolakis, 2005].
5. An integrated healthcare information system

In order to give a better idea on the suggested solution, an exemplar case is presented in the following. The example examines the case of a patient who is visiting a hospital for a routine check.

Upon patient’s arrival at the hospital, his/her Electronic Health Record is retrieved based on the Patient’s Identification Number. The reception sends an HL7 message to all laboratory departments of the hospital that will carry out the check. The results of every test are semantically matched to LOINC as described in section 3.2 and are converted into XML. The results are forwarded for annotation to the doctor who is responsible and are converted into CCR format in order to be incorporated to patient’s EHR as described in section 3.1. The same procedure is repeated for all the medical tests and the patient record is updated.

At the end of the check, the patient record is forwarded to the accounts office and insurance and financial information are updated. The insurance company gets informed of the medical examinations performed and the total cost. The patient gets informed on the tests results by receiving a printable version of the annotated results in an appropriate format. The EHR is updated and is send to the server for filing.

Several issues of concurrent access to the patient’s EHR or of different versions of information stored into distributed information systems have already been confronted in other information systems, so can be easily solved in the Healthcare case.

6. Conclusions

Healthcare services comprise very complicated procedures, with many information systems cooperating for the common wealth. Co-operation demands exchange of mutually understandable information and this requires physical connectivity and common language. If we compare the Healthcare Information Systems to the tower of Babel, the only way to make things work is to build on popular structure standards by exploiting the power of semantics and their matching mechanisms. With the proposed solution, healthcare providers are able to use their own formats, structures and semantics, provided that they make available the appropriate semantic and syntactic matching services. Finally, healthcare information will be available to any authorized system or person and will be delivered in different formats and levels of granularity based on the person that access and the web service that delivers it.

References


Web Resources

[ATC] Official ATC website, hosted by the WHO Collaborating Centre for Drug Statistics Methodology http://www.whocc.no/atcddd

[CRC] E2369-05 Standard Specification for Continuity of Care Record (CCR), Developed by ASTM


[ReadCodes] http://www.equip.ac.uk/readCodes/docs/index.html
