Proceedings of the 7th International Conference on Applied Informatics
Eger, Hungary, January 28–31, 2007. Vol. 1. pp. 335–342.

Object-relational EH databases

Lajos Kollár[a], Henrietta Sipos[b], Krisztián Veréb[a]

[a]Department of Information Technology
Faculty of Informatics
University of Debrecen
e-mail: {kollarl,sparrow}@inf.unideb.hu

[b]Health College Faculty
University of Debrecen
e-mail: sheni@de-efk.hu

Abstract

Nowadays, much effort has been taken to standardize the EHR systems to ease messaging between health databases. In the EHR standard, the model is based on object-oriented technology, and the common language for sending messages is XML. At this point, the need for storing sent and received messages arises. To reach this goal, we have to create a database where such XML messages can be stored and processed, and should give a framework to query them. Several requirements against EHR systems include to be interactive, interoperable, secure and real-time which are also addressed by traditional relational and object-relational databases so it is straightforward to use such a DBMS for data storage. In this paper, we will show some of the various tasks to do about the introduction of the EHR standard and we will show the foundation of an Electronic Health Database based on the features of Oracle 10*g*.

Keywords: Electronic Health Record, XML, Oracle

1. Introduction

The electronic health record (EHR) is (ideally all) patient medical information (including current and historical) from multiple sources, accessible from any location by any provider caring for the patient. An electronic health database is a kind of a multimedia database containing both textual information about the patient's health status and a set of x-ray, CT, MRI, etc. images in EHRs. Pieces of medical information to be transferred from one database to another are called messages. In the last decade, the usage of health databases based on relational database management systems became more and more popular. But there is a big disadvantage of them, namely, that they are not standardized yet and therefore they can hardly

interoperate (with each other and other systems, as well). Hence, the information flow between such health databases is very difficult and laborious.

The previous paragraph can be considered as brief defintions of EHR and the attached terms. In the literature there exists more specific definitions as well. For example, the HIMSS [5] uses the following one: The Electronic Health Record (EHR) is a secure, real-time, point-of-care, patient-centric information resource for clinicians. The EHR aids clinicians decision making by providing access to patient health record information where and when they need it and by incorporating evidence-based decision support. The EHR automates and streamlines the clinician's workflow, closing loops in communication and response that result in delays or gaps in care. The EHR also supports the collection of data for uses other than direct clinical care, such as billing, quality management, outcomes reporting, resource planning, and public health disease surveillance and reporting.

Of course, there are other definitions as well. But all EHR definitions are common in their goals, namely, reducing the paperwork, ensuring the accuracy, eliminating manual charting, eliminating lost or miss-filed reports and finally enhance patient care.

2. EHR standards

One of the first efforts was to create the Computer-based Patient Record (CPR) in 1991, which was published by Institute of Medicine (IOM). It was the first approach which had differentiated the various type of the records. So it talks about EHR or CPR, which is a common terminology for all type of record and PHR (Patient Health Record), which contains information about the patient. The IOM also tekes care to define not only the CPR but the CPR system (with content, format and function) needed to support the EHR. It can be considered as the beginning of the EH databases.

The second bigger effort was the American Society of Testing and Materials (ASTM) EHR E1381-01 in 2001. The ASTM society declared that the EHR is a document (this is a very important declaration, because the most common document transport and storage format is the XML, so it is recommended to use for the electronic health records, as well). This model also differentiates the varios record types, and introduces the longitudinal patient record, which stands for the permanent significant chronological sequenced patient records, containing historical data.

ISO TC 215 Working Group (2002) is also worth mentioning. It has an idea, that an EHR might be virtual. It means, that a virtual EHR consists of some form of logical view or physical assembly of two or more EHR extract from two or more distributed EHR sources (as well as systems). These extracts may belong to a single clinician or healthcare organization but reside on different locations (hospitals, departments etc.). This notion leads to the term of extract and the needs of the various viewing techniques.

Novadays, there are a lot of different EHR standards in the world. It is understandable, because there are lots of different helthcare areas and legacy systems, which meets the different country laws and healthcare habits. Let us see a small outline of them in brief.

The biggest one is the Health Level Seven (HL7) [6]. It was discussed in 2001 in the Standards Insight. The HL7 formally expanded its mission to include support for sharing electronic health records in the USA. It created an EHR Special Interest Group (SIG) to attempt to define the EHR. It has a lot of underlying models (most of them are defined in UML). The full model encompasses not only the small EHR model, but the applications and the infrastructure necessary to deploy an EHR. It is very flexible, hence it is an advantage and a disadvantage simultaneously, because it is too complex and too expensive to apply in each environment. The HL7 uses the term of RIM, which stands for the reference information model.

The openEHR [1] is an international open-source standard. The openEHR Foundation is a non-profit company, limited by guarantee. Its founding stakeholders are University College London, UK and Ocean Informatics Pty Ltd, Australia. The openEHR Foundation is dedicated to the development of an open, interoperable health computing platform, of which a major component is clinically effective and interoperable electronic healthcare records. It does this development by researching clinical requirements, and creating specifications and implementations. The specifications take the form of modular information models, service models and clinical information models. It is very important, that openEHR uses the archetype systems, hence it introduces an archetype parser. (Archetypes are special medical and healthcare rules, they are independent of the applications and the technical systems).

The Europian stnadardization process tried to reach the data interoperability from different starting points. In the early phase some architectural standards have been created relatively shortly, which take efforts to built up hospital systems (HISA) and electronic patient records (EHCRA). They had relevant role in the realization of the KTI program, supported by the IMF at the end of 90's. Unfortunately, the system providers did not follow the predefined standards.

In 1999 the CEN 13606 standard has been released, which tried to standardize the communications using message standards, but the supporting projects started to spread out slowly. The structure of the CEN 13606 looks like openEHR, but till 2004 the whole standard have been refactored, and some bigger regional and international communication applications has deployed.

This resulted in a need for an universal EHR standard.

3. On the "universal" EHR

The universal health record is based on the double model theory. It means that the standard has to contain a reference information model (RIM), and based on this RIM, numerous different variable part (archetypes, etc.) can be applied. The RIM is usually a stable, rarely changing object model (hard part), used by

the application development. It is for the structure. The soft part (often changing part) is the archetype system stands for the content. It formalizes semantical rules for the complex data packages which builds on the RIM entities (e.g. in Hungary, which OEP codes have to be used, etc.).

The suitable modeling tool to create such a RIM is UML. But portability and flexibility is a very important thing in the universal EHR philosophy, that is the reason why a weak UML model has to be used. It means that special UML extensions cannot be used, and the object model has to be able to be implemented in a lot of different languages (Java, C#, etc.). For the interoperability between the systems in a heterogeneous environment, the most suitable communication language is the XML. The interoperability has several levels, as it can be seen in the Figure 1.

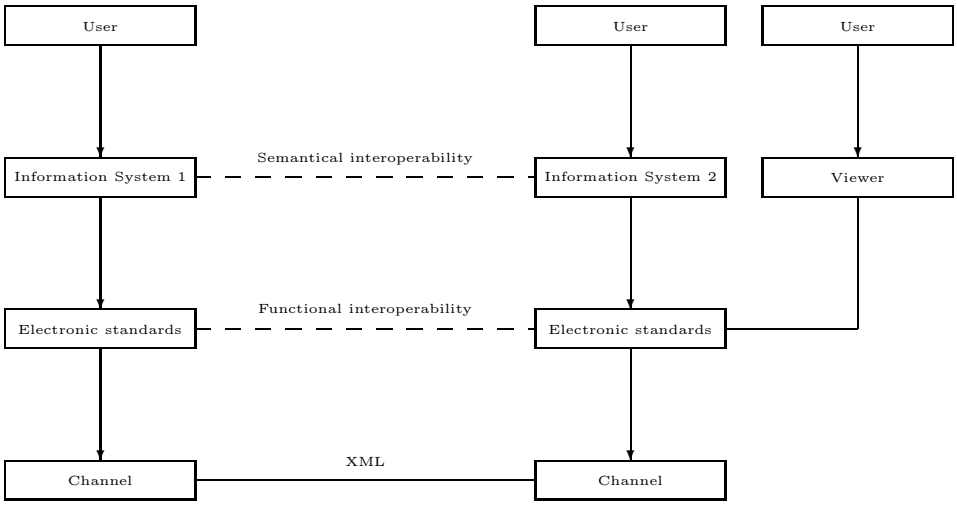

Figure 1: Interoperability levels

It is an important requirement, that the users could communicate each other using the Healthcare Information Systems. This interoperability is a semantical interoperability, where users could see the same information about the same entity (patient, etc.). But systems could differ in a heterogeneous environment, so electronic standards have to be used for the communication. If the two system "talk" in the same "language" i.e. they use the same standards, we could say, that they realize the functional interoperability. To establish that, electronic standards should be independent of the communication channel (TCP/IP, SMTP, etc.) used, and some standard communication language is appropriated to be used, for example the XML.

It is also an important requirement, that the users want to listen the communication without the expensive healthcare systems, so some viewers have to be

implemented, which are able to view the healthcare records as a single-side application.

4. The MSZE 22800

As it was anticipated, the standardization process started to spread out at the Hungarian healthcare communities, as well [2], [3], citeehr3, citeehr4. The standardization efforts resulted in an accepted prestandard. It is called MSZE 22800. It is a quasi "universal" EHR, so it tries to reach the goals of that, i.e. it is built on the double model theory. More precisely, one can classify the prestandard parts into two groups, namely, the RIM standard and the messaging standard. The MSZE 22800 calls the RIM as a HRIM Common Model. It is referenced by number 22800-1. The messaging standards contain more different parts. They are e-Kórlap 22800-2, e-Konzílium 22800-3, e-Lelet 22800-4, e-Recept 22800-5, and finally e-Finanszírozás 22800-6.

HRIM Common Model 22800-1 has an UML model specification. In the class hierarchy the main class is the EHR_Extract (it is not the top class in the hierarchy, but the centre of the messaging model). The instances of them could contain the various versions of the Record instances. The records have various formats (they aggregate Element, Entry or Cluster instances). The basic types (intervals, strings, number values, etc., are stored in the records as attributes, or aggregated complex values. For more details please read the whole specification [7], which can be accessed as an Enterprise Architect project [8].

For grouping the messages in the prestandard 22800, it is a good approach to study the number of patients involved in the message. The "one patient" messages are the e-Kórlap (hospital chart for one patient), e-Konzílium (hospital chart for one patient), e-Lelet (peace of an individual file) and e-Recept (on the available and required medicians and drugs). For more patient there is e-Finanszírozás (electric finance records). For no patient the prestandard enables to create such messages, but it is not part of the 22800. By an other terminology the service messages (messages between the vendors) are the e-Konzílium (request message) and the e-Lelet (answer for the request).

In the international standardization processes it is suggested to use the "one patient" messages. Any other message should be considered as an extentional message. For the extentional messages the prestandard use an EHR_Extract instance which has a a subject_of_care attribute. This attribute means that the message is one patient message. If the record stores information about more patients, the message contains a Report instance, which can contain arbitrary number of EHR_Extract instances. If there is no patient concerned to the message, there must be one or more Any_Extract instances to store such data.

5. EH Databases

The existance of medical record implies the need for storing, managing and retrieving them. An electronic health (EH) database is a kind of a multimedia database containing both textual information about the patient's health status and a set of x-ray, CT, MRI, etc. images in EHRs. In our notion, an EH database more than a database containing the patients' data. It has to support messaging and the underlying RIM as well.

Lot of database management system can handle RIM data. The main healthcare system provider is Oracle. The Oracle Healthcare Transaction Base (HTB) [4] is based on the E-Business (Oracle Application) platform and HL7. Oracle HTB has two layers. The first is the HTB API (Application Programming Interface) and the second is the HTB core. The former is a collection of function and procedure specifications, and the latter is a set of stored procedures, types, packages, objects and collections to handle the physical data. In this case, there is no application layer, it is just a "medical utility" for the Oracle-based healthcare systems.

The Oracle HTB is used by the Oracle Healthcare Intelligence (OHI) [12], which is built on top of HTB. It supports a variety of industry reporting needs. It contains a loader program, schematas and applications. The loader can extract data from the HTB, and the OHI populates them to the schematas, and the application calculates the facts as needed. The OHI could be a basis layer of a higher level healthcare application.

The Oracle HTB has several components. They are the Enterprise Object Model (which is formerly the HL7 RIM), the Enterprise Terminology Services (which supports multiple type of terminologies for the different countries), the Clinical Services (which is the main application component) and others for handling the application (e.g. Security Services, Configuration, etc.).

The power of the HTB can be found in its functionality. That first one is the Administrative and Clinical Business Services, where one can handle the enterprise master data, and it can help us in the organization management, documentation, etc. The second main functionality is the Core Application Services, where the messaging, the terminology and the security can be handled. In the Domain-specific Architecture the local patient records versioning can be served, and finally, there is the Business Intelligence functionality, which is for the higher level applications.

But Oracle HTB has some disadvantages as well. The main problem comes from its complexity. It is too difficult to introduce for every hospital, needs a lot of training and too expensive (mainly in Hungary). And, of course, it does not support MSZE 22800, because it is built on HL7.

To solve this problem we work on such a system, which can store, process and retrieve Hungarian MSZE 22800 messages and data on object-relational basis. In our case, the main ORDBMS is Oracle [11]. In our solution, original XML messages can be stored, processed and searched via standard tools (with the Oracle built-in XML parser invoked by stored procedures). Records can be searched, sorted and listed via SQL statements, as well. Another huge advantage is that medical images

(stored as BLOBs and Still Images or ORD Images) can be processed and displayed via *inter*Media (where a lot of image processing tool can be used with stored java procedures). The layers of our system can be seen on Figure 2.

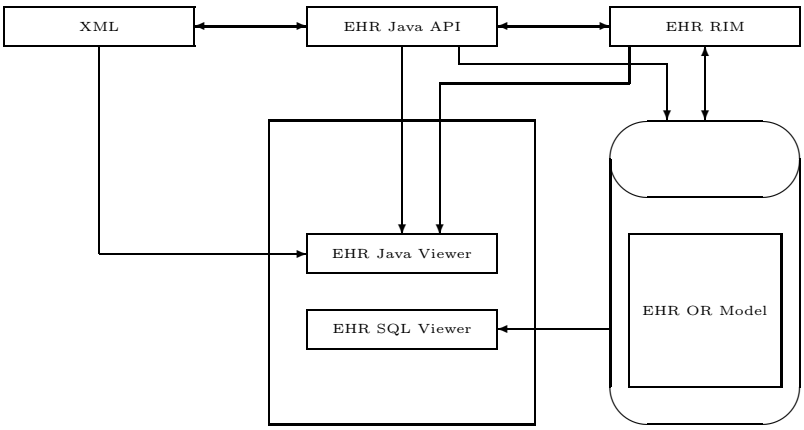

Figure 2: The layers of our system

The EHR Java viewer can process XML messages directly, or using the standard EHR Java API. The EHR Java API transforms XML messages into EHR RIM instances. XML messages can be stored in the EHR Object Relational Database in an OR schema, so the whole EHR RIM is mirrored as an EHR OR data model. The database uses a Java API for storing and retrieving data from the RIM. The SQL viewer accesses the stored EHR data by direct SQL and PL/SQL invocations.

6. Future work

The above mentioned work has successfully been started, so the figured layered system was started to developed. The EHR Java viewer has already been developed as a direct XML viewer, it does not use the EHR Java API yet, because the API is under development. The indirect part of the Java viewer is being developed. The main part, the OR data model of the EHR RIM is ongoing, in line with development of the Java API. Finally the SQL viewer should be developed.

Our goal is to deploy a well-useable system for serving the needs of a future healthcare system, implementing the MSZE 22800, providing a large-scale of functionality and robostusness of the underlying Oracle Object-Relational Database Management System.

References

[1] Beale, T., Heard, S., Kalra, D., Lloyd, D., The openEHR EHR Reference Model, www.openehr.org

[2] Horváth, L., Puskás, Zs. P., Héja, G., Nagy, I., A hazai egészségügyi elektronikus adatcsere szabványainak adatmodellje, *Infokommunikáció-eEgészségügy*, 4(1), (2005), 40–46.

[3] Horváth, L., Puskás, Zs. P., Héja, G., Nagy, I., A hazai egészségügyi elektronikus adatcsere üzenetszabványai, *Infokommunikáció-eEgészségügy*, 4(2), (2005), 45–47.

[4] Healthcare Transaction Base, http://www.oracle.com/industries/healthcare/healthcare-transaction-base-datasheet.pdf

[5] HIMSS, An analysis of Health Information Standards Development Initiatives, April 2003, HIMSS Standards Insight.

[6] HL7 V3 Reference Information Model, V02-02, www.hl7.org.

[7] MSZE 22800 Magyar Előszabvány, Egészségügyi Informatika, Magyar Szabványügyi Testület, (2004).

[8] MSZE 22800 Magyar Előszabvány Enterprise Architect UML model, http://www.euagazat.hu/portal/server.pt/gateway/PTARGS_0_237_3225_0_0_18/index

[9] Puskás, Zs. P., Horváth, L., Héja, G., Nagy, I., eSzabványok: Kommunikáció szabványos üzenetekkel, *Infokommunikáció-eEgészségügy*, 4(3), (2005), 40–42.

[10] Puskás, Zs. P., Horváth, L., Héja, G., Nagy, I., eSzabványok: Üzenetek, dokumentumok, adatok, *Infokommunikáció-eEgészségügy*, 4(4), 2005, 41–43.

[11] Oracle Database 10g Enterprise Edition, http://www.oracle.com/technology/products/database/oracle10g/pdf/DS_General_Oracleq_Database10gR2_EE_0605.pdf

[12] Oracle Healthcare Intelligence, http://www.oracle.com/industries/healthcare/healthcare-intelligence-datasheet.pdf

Lajos Kollár, Krisztián Veréb
H–4032 Debrecen, Egyetem tér 1.
Hungary

Henrietta Sipos
H-4400 Nyíregyháza, Sóstói út 2–4.
Hungary