**Communication Problems Among Heterogenous EHR Systems**

Nagy, Miroslav
2007

# Communication Problems Among Heterogenous EHR Systems

*Post-Graduate Student:*
MGR. MIROSLAV NAGY

Department of Medical Informatics
Instutite of Computer Science of the ASCR, v. v. i.
Pod Vodárenskou věží 2

182 07 Prague 8, CZ

nagy@euromise.cz

*Supervisor:*
RNDR. ANTONÍN ŘÍHA, CSC.

Department of Medical Informatics
Instutite of Computer Science of the ASCR, v. v. i.
Pod Vodárenskou věží 2

182 07 Prague 8, CZ

riha@euromise.cz

Field of Study:
Biomedical Informatics

This work was supported by the project number 1ET200300413 of the Academy of Sciences of the Czech Republic.

**Abstract**

This paper describes the problem of medical information transportation among heterogenous Electronic Health Record (EHR) systems. First we describe the deatils of the problematics and mention some examples. After that we show a proposed solution methodology and a pilot application of an EHR system called AdamekJ capable of interchanging data via HL7 message standard. Finally some future work will be presented in the field of EHR systems' models comparison and synchronization possibilities of these models.

## 1. Introduction

Nowadays the EHR is becoming an integral part of patients' health documentation. On the basis of computerized form of health information it is possible to personalize the health care and make better use of medical knowledge and guidelines. However, there are difficulties applying traditional aproaches in the field of information system development, the best results in standardization and computer science must be employed.

Since in the praxis it is very difficult to abandon current systems or modify them completely, the openness and modularity of used systems enabling integration of heterogenous medical data originating from different sources becomes crucial. To create such a distributed healthcare environment, where medical information, is commonly shared the use of communication standards is inevitable. One of such standards is e.g. HL7, which has its origin in U.S.A. European standard CEN EN 13606 deals with EHR architecture as well as interoperability via messaging. In the rest of the paper we consider messaging services among EHR systems rather than the methodologies of building of EHR systems. Since the HL7 standard dominates in the field of communication, we aim our interest to it.

In the Czech Republic there are many commercial medical information systems on the market. Most of them do not concern the storage of medical data in a form that would conform the definition of EHR according to [1]. These systems are mainly used to manage the health services, financing and ensure the functioning of the whole health providing facility. However, storing information about patient's health in computerized form is gradually getting into the center of the interest.

## 2. Motivation

One could ask why we need a distributed healthcare. As the European Union accepted new members, people started to migrate in a bigger scale than it was in the past, and therefore the importance of interoperable access and integration of the distributed information arose. Another field of application of distributed medical information is the case of an emergency rescue, when every information about patient's health or treatment is of vital importance.

Since there exist more than one communication standard, the first step in simplifying the data exchange process is to use the same standard on both sides of the communication channel. If different medical standards are used, it is necessary to map and transform the messages to each other.

Usually, the greatest problem is to agree on the same standard. This situation prevails because there is no universally accepted standard for the electronic representation of clinical data. One of many reasons of disagreement is the economical and political background, since governments invested huge resources into research and the development of their national standards.

When two parties reach an agreement and healthcare providers communicate the same standard, it is usually a country dependent standard and thus new problems arise when its exposed to international use. For example in the Czech Republic, there was developed a data standard named DASTA [2]. This standard was designed for laboratory results interchange in the first place. In the course of time the range of structured data broadened but it still does not cover most concepts in medical domain. However, it is under continual development, it is still incompatible with other EHR standards, thus unsuitable for application in international scope.

For example, the project called ARTEMIS [3] dealt with interoperability problems among medical information systems storing clinical information in various proprietary formats.

The definitions of the terms "archetype" and "template" are necessary for proper understanging of the further text. In the paper [4] they are stated as follows:

- *archetype* – a computable expression of a domain content model in the form of structured constraint statements, based on some reference model. openEHR [5] archetypes are based on the openEHR reference model. Archetypes are all expressed in the same formalism. In general, they are defined for wide re-use, however, they can be specialized to include local particularities. They can accommodate any number of natural languages and terminologies.
- *template* – a directly, locally usable definition which composes archetypes into a larger structure logically corresponding to a screen form. A templates may add further local constraints on the archetypes it mentions, including removing or mandating optional sections, and may define default values.

Some examples of archetypes can be found at [6]. In the Figure 1, the structure of archetype representing blood pressure concept is depicted. The part *data* contains values of the actual pressure, i.e. systolic, diastolic, mean arterial pressure, pulse pressure and textual comment on blood pressure reading. *State* is a list of information describing conditions of the meassurement, e.g. the position of the patient at the time of meassuring blood pressure. *History* covers separate meassurements and adds temporal data in the implicit form, i.e. base meassurement in the history, another reading after 5 minutes of rest, 10 minutes etc. Finally, the *protocol* holds technical data such as size of a sphygmomanometer's cuff if it is used or a specification of an instrument used to measure the blood pressure. For the sake of further computerized processing, archetypes are defined in ADL (Archetype Definition Language) [7].

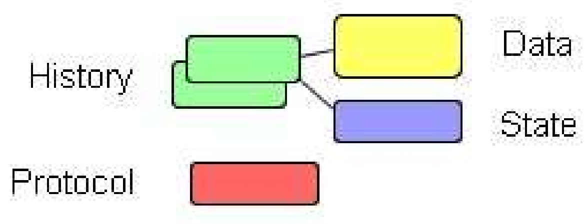


**Figure 1:** Structure of blood pressure archetype (id openEHR-EHR-OBSERVATION.blood_pressure.v1)

The solution of the ARTEMIS project contained an idea of wrapping and exposing the existing healthcare applications as Web Services [8]. The semantic interoperability was achieved by using OWL [9] (Web Ontology Language) mappings of archetypes based on reference models of, possibly, different standards (e.g. openEHR, HL7 RIM). These archetypes semanticaly enrich the Web Services messages. The interoperability was realized through a mediator that transformed the source message using mapping definitions into appropriate form to be accepted by the destination system and its Web Service.

## 3. Using heterogenous models

Another problem in sharing medical data is the possibility of different definitions of concepts despite using the same modeling methodology (the term modeling methodology comprises all steps necessary to create the resulting model).

Heterogenity in models occurs when there is a disagreement about the meaning, interpretation or intended use of the same or related data. Usually two separated individuals (experts, developers etc.) model the same domain in more or less different way, even when using the same methodology.

The similarity or heterogenity of the models can occur on two levels. The first one is *the functional* (implementational) level where information systems use to communicate diferent network protocols (e.g. IP - Internet Protocol), transport binding (e.g. HTTP, FTP) or message format (e.g. XML, ASCII text). The second level is *the semantical* where systems have to understand each other's formal definitions of domain concepts. The latter will be the one of our concern.

### 3.1. MUDR EHR and WinMedicalc 2000 comparison

In the project "Information technologies for development of continuous shared healthcare" (no. 1ET200300413), we deal with the problem of sharing medical information among EHR system developed by various vendors. To fulfil the project's objectives we decided to implement the medical data exchange between two particular EHR systems using HL7 v.3 standard. One of used EHR systems is the MUDR EHR developed at the EuroMISE Center in the Institute of Computer Science of the Academy of Science of the Czech Republic and the second is a commercial application called WinMedicalc 2000 created by Medicalc Software s.r.o.

MUDR EHR uses so called knowledge trees to model stored information [10]. The WinMedicalc 2000 stores its data in a relational database and thus uses Entity-Relationship model [11] to represent its information model. Development of both EHR's started from the same modeling basis. Each originaly used the so called minimal data model of cardiology [12].

In MUDR EHR, the modeling process resulted in creatiing of a knowledge domain called `PATIENT`, consisting of basic administrative data, allergy information, family history, social history, subjective information, physical examination, lab examination, personal history, treatment information and history of cardio-vascular diseases.

The model of WinMedicalc 2000 system consists of basic administrative information, cardiological examinations (e.g. ECG examination, Holter monitor, stress test ECG etc.), lab examination, physical examination and family history. Each of these data (except administrative information) are connected to a clinical event, that binds together the object and subject of the event, i.e. the patient and the physician. Clinical event that contains yet another information such as place where the event took place (e.g. ward, emergency room).

Moreover, WinMedicalc 2000 system covers a broader scope than just clinical data (e.g. catering services, bed management), but these are out of our concern so we leave them out. To illustrate the simmilarity and difference between MUDR EHR and WinMedicalc 2000 system, the Figure 2 shows the screenshots of client applications.

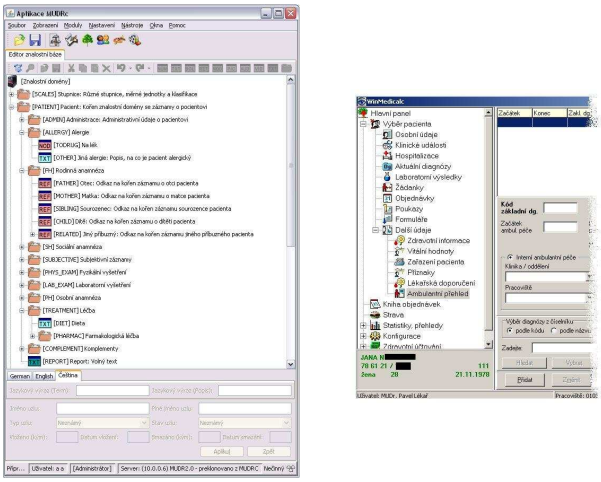

**Figure 2:** MUDR client, application MUDRc, and WinMedicalc GUI

### 3.2. Message interchange

The result of solving partial tasks in the project "Information technologies for development of continuous shared healthcare" is a proposal of communication schema between MUDR EHR and WinMedicalc 2000 system using HL7 messaging standard. In these days the communication schema is being implemented and partially tested. The communication between these two health information systems (HIS) can be divided into following steps:

1. **HIS1** retrieves the required data from its repository.
2. Retrieved data are written into the template based on R-MIM that models the content of the message and origins in the information model of the **HIS1**. HL7 template filled with data is sent to the **Translator1**.
3. **Translator1** transforms the template filled with data into HL7 message and sends it back to **HIS1**.
4. The received HL7 message is sent through the network to the receiver.
5. The receiver, the **HIS2**, gets the HL7 message. Since the data is still in the form that **HIS2** does not understand it is posted to the **Translator2**.
6. **HIS2** gets back a template that is derived from R-MIM of **HIS2** and is filled with data that correspond to the structure of desired message.
7. Finally the required data is stored to the repository of **HIS2**.

The algorithm mentioned before is graphically represented in the Figure 3.

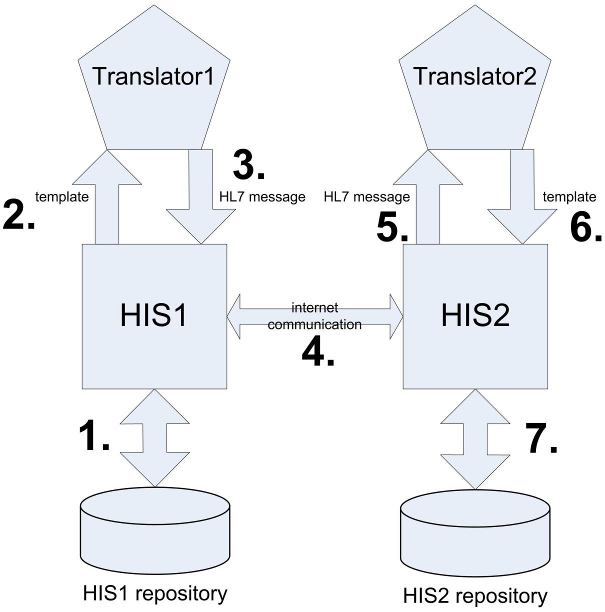


**Figure 3:** Communication algorithm between MUDR EHR and WinMedicalc 2000

### 4. Models' comparison on semantical level

Schemagic [13], a synchronizing tool formerly designed to compare and synchronize database schemas stored in relational databases, will be extended by a plug-in capable of processing archetypes. The input of the plug-in are two archetypes described in ADL (Archetype Definition Language) [7]. The extended tool would be helpful in finding differences between two given archetypes describing the same concept but originating from different sources, e.g. two professionals or two EHR systems developed by different vendors.

The rough version of methodology used to establish communication between two EHR systems is described in the following steps:

1. Take models formalizing EHR systems which want to communicate.
2. If chosen models were created using different modeling techniques, transform one model into the same form as the other and go to the next step, else go to next step immediately.
3. Use some comparing tool to find differences between models.
4. To remove differences alter models in appropriate way.
5. Create transforming modules for messages corresponding to both models.
6. The communication can be established.

The first step assumes that both EHR systems do have a formal description in a form of a model. If one or both miss such a model, it is necessary to create one, using the same procedure, i.e. methodology if possible. The step number 2 checks if the same modelling methodology (the same standard or formalism) is used. If not, an extra transformation is needed. This transformation covers modifications on syntactical level. The result of step 2 should be based on the same grounds. Semantical modifications will be discussed later on. The Schemagic extension mentioned above will be helpfull in step 3. Alterations mentioned in step 4 will be accomplished manually, since no model manipulating langugage (such as data definition language – DDL in SQL) in the domain of EHR standards exists.

The next planned extension of the Schemagic would ease the EHR interoperability by simplifying the development of the EHR system modules implementing communication, data exchange, based on some standard such as HL7. This extension is bound in step 5 of the methodology mentioned above. Our pilot implementation will take into account only HL7 version 3 standard. The task of the Schemagic's extension is to map HL7 templates derived from R-MIM model with HL7 message fragments called CMETs (Common Message Element Types). Using this tool will result in much shorter implementation time of translator module (see **Translator1/2** in Figure 3) of EHR system translating specific HL7 template into HL7 message.

### 5. Pilot application implementation

A pilot application has been developed to test ideas discussed above. Its name is AdamekJ and comes from an abbreviation of ADAMEK (stands for "Aplikace Datoveho Modelu EuroMISE-Kardio" – Application of Data-Model EuroMISE Cardio) [14]. The letter 'J' stands for Java, since AdamekJ is a Java application.

We test communication between WinMedicalc 2000 system and AdamekJ rather than MUDR EHR, because MUDR EHR contains only testing data that are sufficient for determination of HL7 standard usability in the Czech environment. The AdamekJ application will be deployed in the ambulance of preventive cardiology, thus it will contain real "production data" which can more precisely show convergence or divergence of clinical content of communicated messages.

### 5.1. AdamekJ

Application ADAMEK [14] was developed to collect data in the ambulance of preventive cardiology run under Institute of Computer Science, Academy of Sciences of the Czech Republic. It was created in 2002 and is still in use. Since the application was implemented as a standalone MS Access 2000 program, it reached the limits of the used database. As soon as suitable tools to implement a more advanced version were available, we started the development of the AdamekJ application as the successor of ADAMEK. Both applications, ADAMEK and AdamekJ, are based on the minimal data model of cardiologic patient [12].

During the design process an indispensable emphasis was layed on usage of modern technologies. AdamekJ is a two-layer application consisting of the data layer and the user interface. Application domain objects are persistently stored into relational database Oracle 10g [15]. Objects' persistency is achieved by using the Hibernate framework [16]. The framework is configured by XML mapping files (HBM – Hibernate Mapping). HBM files map objects from application's domain object model to relational tables in the database.

Core classes of the application are implemented using Spring Framework [17]. Spring is a layered Java/J2EE [18] application framework. J2EE (Java 2 platform, Enterprise Edition) is the industry standard for developing portable, robust, scalable and secure server-side Java applications. Building on the solid foundation of the Java Platform, Standard Edition (Java SE), Java EE provides web services, component model, management, and communications APIs that make it the industry standard for implementing enterprise-class service-oriented architecture (SOA) and next-generation web applications. Spring Framework provides automated configuration and wiring of application objects. Spring is well integrated with Hibernate and simplifies the configuration of domain objects' persistent storage.

User interface is implemented using Spring RCP (Rich Client Project). Spring RCP is based on Java Swing, thus the resulting application is a Swing application. The main advantage of the Spring RCP project is providing an elegant way to build highly-configurable, GUI-standards-following rich-client applications faster by leveraging the Spring Framework, and a rich library of UI factories and support classes.

The AdamekJ application is in its testing phase and being prepared for deployment. The screenshot in the Figure 4 shows the detailed view on physical examination of a patient.

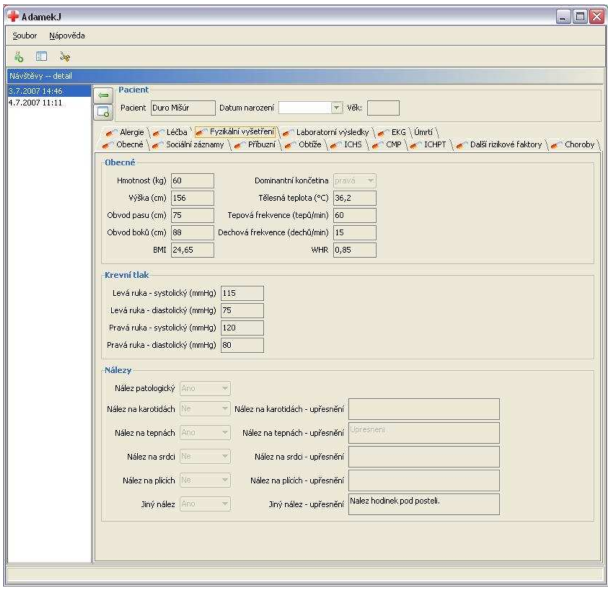


**Figure 4:** User interface of AdamekJ application

An integral part of each EHR is its communication with other systems in health care environment. Some systems are just limited to import and export data, but AdamekJ stands before the phase of implementation of HL7 messaging standard. When this phase is finished, we will be able to evaluate the communication in the form of HL7 messages between two heterogenous EHRs based on minimal data model of cardiological patient.

## 6. Conclusion

The first step towards fulfilling the goal of interoperable EHRs is making the implementation of messaging standards easier. We proposed an extension of the system Schemagic that would find appropriate balloted HL7 message fragments corresponding to clinical content of given EHR. This will result in significant reduction of time needed to develop a transforming module, i.e. translator depicted in Figure 3.

Next, the archetype comparison and harmonization will be studied and implemented as further extension of the Schemagic system. However, there still remains the difficulty how to decide whether two archetypes model the same concept or not. This problem will be under further systematic exploration.

## References

[1] P. Schloeffel, "Electronic Health Record Definition, Scope, and Context", *ISO/TC 215 Discussion Paper*, 2002.

[2] Ministry of Health of the Czech Republic - MZ ČR, "Datový standard MZ ČR v.4" *Ministry of Health of the Czech Republic*, http://ciselniky.dasta.mzcr.cz, (last accessed on 8.7.2007).

[3] ARTEMIS Consortium, "The Artemis Project Home Page", http://www.srdc.metu.edu.tr/webpage/projects/ artemis, 2004.

[4] T. Beale, S. Heard, "Archetype Definitions and Principles", *The openEHR foundation*, Rev. 0.6, 2006.

[5] The openEHR Foundation, "openEHR Homepage", *CHIME at University College London*, http://www.openehr.org/, (last accessed on 8.7.2007).

[6] Ocean Informatics, "Archetype Finder" *Health Informatics, Central Queensland University and Ocean Informatics Pty. Ltd.*, http://www.archetypes.com.au/archetypefinder/ archetypefinder, (last accessed on 8.7.2007).

[7] T. Beale, S. Heard, "Archetype Definition Language (ADL)", *The openEHR foundation*, Rev. 1.3.1, 2004.

[8] V. Bicer, O. Kilic, A. Dogac, G. B. Laleci, "Archetype-Based Semantic Interoperability of Web Service Messages in the Health Care Domain", *Int'l Journal of Semantic Web & Information Systems*, 1(4), pp. 1–22, 2005.

[9] Ed.: D. L. McGuinness, F. van Harmelen, "OWL Web Ontology Language", *W3C Recommendation*, http://www.w3.org/TR/owl-features/, 2004.

[10] J. Spidlen, "Databázová reprezentace medicínských informací a lékařských doporučení" (In Czech), *Master Thesis at Faculty of Mathematics and Physics, Charles University in Prague*, pp. 32–34, 2002.

[11] P. Chen, "The Entity-Relationship Model – Toward a Unified View of Data", *ACM Transactions on Database Systems*, Vol. 1, No. 1, pp. 9–36, 1976.

[12] M. Tomeckova, "Minimální datový model kardiologického pacienta" (In Czech), *Cor et Vasa*, Vol. 44, No. 4, pp. 123, 2002.

[13] M. Nagy, "Synchronization of Relational Schemas", *Master Thesis at Faculty of Mathematics and Physics, Charles University in Prague*, 2005.

[14] R. Mares, M. Tomeckova, J. Peleska, P. Hanzlicek, J. Zvarova, "Uživatelská rozhraní pacientských databázových systémů – ukázka aplikace pro sběr dat v rámci minimálního datového modelu kardiologického pacienta" (In Czech), *Cor et Vasa*, Vol.44, No. 4, pp. 76, 2002.

[15] K. Loney, B. Bryla, "Mistrovství v Oracle Database 10g" (In Czech), *Computer Press*, 2006.

[16] Ch. Bauer, G. King, "Hibernate in Action", *Manning*, 2004.

[17] C. Walls, R. Breidenbach, "Spring in Action", *Manning*, 2005.

[18] Sun Corporation, "Java EE at a Glance", *Sun Corp.*, http://java.sun.com/javaee/, (last accessed on 8.7.2007).