[2] European Committee for Standardization (CEN), Technical Committee CEN/TC 251: prEN 13606-1, Health informatics – Electronic healthcare record communication – Part 1: Reference model (2003)
[3] Špidlen Josef: Databázová reprezentace medicínských informací a lékařských doporučení (in Czech), Master's Thesis, Faculty of Mathematics and Physics, Charles University in Prague (2002)
[4] Špidlen Josef: Electronic Health Record Communication XML Schema [http://www.euromise.cz/EHR/MUDRAPI.xsd]

Acknowledgments
The work was supported by the grant number 1ET200300413 of the Academy of Sciences of the Czech Republic.

Address
Mgr. Ján Ďurovec
Department of Medical Informatics,
Institute of Comupter Science AS CR
Pod Vodárenskou věží 2
182 07 Prague 8
e-mail: durovec@euromise.cz

Interoperable Electronic Health Records in Continuous Shared Health Care

Hanzlíček Petr, Ďurovec Ján, Nagy Miroslav, Zvárová Jana, Špidlen Josef
Department of Medical Informatics, Institute of Computer Science AS CR, Prague, Czech Republic

Abstract
Changes in present health care provision require higher use of information and communication technologies to support the cooperation during the treatment. The project *"Information Technologies for Development of Continuous Shared Health Care"* deals with the development of the methods and technologies helping to ensure the continuous shared health care in information society by creation of the unified system of electronic health documentation including the use of electronic identification cards and emphasized security aspects. The technological solution will be based on the results of previous research and development in the EuroMISE Centre, results of the European research and international standards of health informatics.

Introduction

In the last decade many social, political and economical changes in many European countries influenced the development of traditional models of social care and health care. The nature of these changes (demographic changes, increased costs of health care, need for improvement of health care quality, etc.) demands changes in present health care provision and higher use of information and communication technologies. These activities require storage, sharing and exchange of data. When we think about lifelong healthcare documentation, we see a lot of different groups of healthcare professionals taking care of a patient, performing different tasks. All these groups need to share the information to be able to cooperate during the treatment. The need to limit the expenses of healthcare also demands a new structure of health care and change of task of the primary care, public care and home care. All these changes should be reflected by the coming information society in healthcare.

Healthcare documentation contains both the information related to the physical or mental health of the patients and the information related to the provision of health care by healthcare professionals or institutions. The personal healthcare information provided by the patient can be extended by his relatives, by social workers or by other persons. The provision of health care is stored into professional records containing examinations and opinions of physicians and other healthcare personnel. These data are entered not only by physicians taking care of the patient, but also by other persons providing additional examinations (pharmacists, pathologists, CT operators…), administrative persons and even patients themselves. This diversity of sources brings a lot of problems concerning the security of the information in the paper-based records as well as the electronic records. The electronic health record therefore has to be created, implemented and operated in a way that limits the possibility to harm the patient.

The electronic health record offers the possibility to formalize the structure of the classical health record, which can bring a more transparent way of recording of healthcare procedures and an immediate access to the health data of the patient. To achieve this, the specialists must agree on the clinical terminology, the definitions and a health record structure. They also have to accept a more strict approach to the storage of healthcare documentation.

The legislative development in the Czech Republic brought several legislative acts, modifying the legislative point of view on the electronic health records. The first one was the legislative act 101/2000 Coll. on the personal data protection, the second one was the act 227/2000 Coll. on the electronic signature. In the year 2001 the novel of the act 20/1966 Coll. on health care was approved. All these legislative acts give a legal frame for keeping an electronic form of healthcare documentation without the need to keep its paper-based form. This situation leads to the expectation of a quick development of electronic health documentation, telemedicine and intelligent systems in Czech healthcare [1], [2].

The applied research in the field of electronic health record is an important tool to connect the healthcare institutions in order to collect and share healthcare information and to guarantee the continual shared health care. The availability of expensive technological and personal sources in health care, independent of the place of these sources, is a clear demand of present medicine. Telemedicine is concerned with many important aspects of the new model of health care delivery and can play an important role in innovation of present medicine using new information and communication technologies. Telemedicine can also change the classical form of health care providing the effective solutions in an increasing number of new situations.

Aims of the project

The Department of Medical Informatics of the Institute of Computer Science AS CR is the main contractor of the project *"Information Technologies for Development of Continuous Shared Health Care"* of the Academy of Sciences of the Czech Republic. The project consortium further consists of two companies specialized in the hospital information system development and in Internet technologies, information systems and research and development in the biomedical informatics area. The project deals with the development of the methods and technologies helping to ensure the continuous shared health care in information society. The research is based on development of methods and techniques of a remote access to the information in a form of data and knowledge in the relationship with electronic health documentation. The basic prerequisite for telemedicine applications is the electronic health record. The telemedicine applications

enable utilization of the information, stored in electronic health record, simultaneously to many persons in different institutions and enable the immediate access to the information related to the current state of the patient. In the frame of the project methods and techniques are being developed, that support the creation of the unified system of electronic health documentation including the use of electronic identification cards. A related research task is the research of methods on the delivery of biomedical data and knowledge to authorized users in a way conformant to the Czech legislature and development of statistical methods for uncertain information processing facilitating the interpretation of collected data. Another research task is the research of GRID technology services in relation to the development of intelligent systems for prevention, diagnostics and care.

Methods

One of the main research tasks is the design, implementation and usability verification of a shared electronic health record as a tool for electronic health documentation and continuous shared health care. The essential topic from the research point of view is the comparison of current technologies of structuring and storing of the data, their security and safety in the context of Czech healthcare and legislature and evaluation of usability of chosen technology as a data source for research purposes (statistics, data mining). The important practical consequence of the project could be the improvement of quality of care and reduction of costs by reduction of a number of repeated examinations and procedures. The electronic cards will be used as an identifier of healthcare professionals as well as patients. The patient cards will carry the additional basic information for emergency purposes. The key part of the proposed solution will be the analysis and implementation of different methods of access control and data protection in the shared electronic health record, possibilities of keeping the encrypted information by third party or server without possibility of non-authorized access to this information etc.

The technological solution is based on the development results in the field of electronic health record in the EuroMISE Centre [3], [4], [5], [6], on the results of European research and on the international standards of health informatics. The concept of a lifelong shared electronic health documentation enabling communication between many different healthcare institutions, healthcare professionals and other organizations cooperating during the process of treatment has some consequences to the architecture of the communication system, the structure of the exchanged information and the safety and security issues.

In order to use the health documentation of a patient in an intelligent way, allowing to implement different decision support systems, to control the access to different parts of the health record for the authorized users or to be able to automatically provide the appropriate information on request from other cooperating institution, the structured electronic health documentation is a key requirement. Free text information, well known from discharge letters provided by specialists after an examination, often consists of vague formulations, ambiguous statements without syntactical correctness and is therefore hardly processible by automatic systems, able to draw attention to possible problems in provided healthcare, e.g. drug contraindications or discrepancy between a treatment procedure and recommendation in medical guidelines. Therefore two electronic health record systems, based on structured storage of the information were chosen for the pilot project, enabling to use some of the above described techniques of decision support of the physician. To ease the extraction of structured data from medical texts (discharge letters, free text reports, etc.), methods of natural language processing were studied, the clinical terminology was analyzed and tools for a free text analysis were developed [7].

MUDR Electronic Health Record

The first of the two EHRs comes from the results of the previous research in the field of electronic health record in the EuroMISE Centre and it is called “MUltimedia Distributed Record” (MUDR). To be able to use a universal technique for processing and visualisation of stored data, a unified way of structured data storage was needed. The set of collectable attributes varies in different departments, organizations and also during time. Therefore we needed dynamically extensible and modifiable structure of items allowing reorganization without a change of the database structure. For information storage we use two main structures described by the graph theory expressions. The set of attributes and their relations is described by a graph structure that consists of vertices representing symptoms and edges representing relations among symptoms. The basic structure underneath is the oriented graph (tree) representing the hierarchy of medical knowledge stored in the patient record by the edge type "subjection". Other types of edges are used to express relations such as equivalence of some terms in different parts of a knowledge tree. We call this tree structure "knowledge base". Each vertex describes the represented attribute by its identification, internal name, type and administrative data like identifications of a user who created or deleted the vertex. The edges between vertices are described similarly. This enables the system to keep a track of any change in the structure. If the structure is changed for any reason, the "deleted" vertices and edges are only marked as deleted and identification of the user who deletes them is stored.

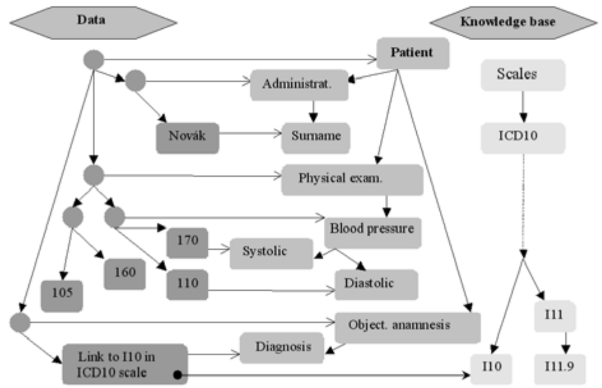


Fig. 2. The illustration of a structure of the knowledge base and patient data in the MUDR EHR.

The collected data are stored using a similar structure as the knowledge base. Information about one patient corresponds to right one tree in the forest. Each vertex in the tree describes one instance of the attribute from the knowledge base by the internal name of the attribute, its value and administrative data. The values are physically stored in separate tables according to their types and are related to their descriptions in the tree.

The MUDR electronic health record (EHR) is based on a three-layer architecture with a data layer, an application layer and a user interface. This decomposition enables to separate different system modules to small functional parts, which makes the system more flexible. The global architecture with communication interfaces based predominately on XML and HTTP has been defined. The interface for connecting the decision support tool implementing the medical guidelines formalized in the form of dynamic libraries was also defined. Based on these experiences, the lightened version of EHR named MUDR-Lite was created. The system uses classical database modelling techniques, the user interface is realized as the interpreter of a special XML based language (MUDR-Lite Language – MLL), describing the individual components of user interface, their states and relations among them.

The other EHR system in the project is part of a commercial hospital information system, practically used in several Czech hospitals. Its architecture is based on a concept of a clinical event, which represents certain identifiable event (examination, surgical procedure, laboratory examination, biopsy, etc.) in the process of patient treatment. The clinical event is mostly associated with some set of data, related to the medical procedure, represented by the clinical event. These data are stored in the central database, structured according to the needs of individual modules of the system.

Solutions

To be able to establish the communication structure, realizing the lifelong shared electronic health record, the clinical content has to be synchronized among the project partners on a broad level. In the pilot project, the large healthcare domain was reduced to the field of cardiology. As the common data set, the minimal data model for cardiology [8] has been taken.

The other task is the selection of an appropriate communication protocol, allowing the exchange of the structured data among partners. After comprehensive research and study of existing and proposed national and international solutions, experience from other countries, international and European standards, we have decided to use the HL7 version 3 standard. Health Level Seven (HL7) is one of several American National Standards Institute (ANSI) – accredited Standards Developing Organizations (SDOs) operating in the healthcare arena. The Health Level Seven's domain is clinical and administrative data. The mission of HL7 is "To provide standards for the exchange, management and integration of data that support clinical patient care and the management, delivery and evaluation of healthcare services. Specifically, to create flexible, cost effective approaches, standards, guidelines, methodologies, and related services for interoperability between healthcare information systems." The HL7 version 2 is widely used worldwide as a communication standard for the healthcare domain, mostly for the communication inside the healthcare institution.

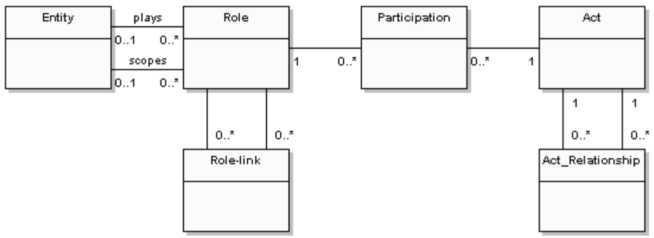


Fig. 3. The six core classes of RIM and their associations.

The HL7 version 3 is a major improvement of the previous standard using the well-defined methodology based on a reference information (i.e. data) model (RIM), which provides an explicit representation of the semantic and lexical connections that exist between the information carried in the fields of HL7 messages.

The first step on the way to achieving interoperability was the creation of a model more specific than the general RIM. A so called LIM (Local Information Model) was created. This type of model is not mentioned in the HL7 version 3 standard and represents RIM connected with vocabulary of

allowed data values. All information systems should include LIM information, as this model in fact describes the information system itself and enables the information system to communicate with other parties. The information system should be also able to process at once all the information that is provided by its manufacturer in LIM templates included in the system.

In the process of communicating the data using the HL7 version 3 standard, the data are being mapped to the general RIM provided by HL7 itself when entering or leaving the information system.

For testing purposes, a sample LIM based on the model of minimal cardiological health record data has been developed. Based on this information model the common templates have also been created. These templates link the building blocks specified by the LIM and define the structure of a particular piece of information to be transferred. However, the data types are not mentioned at this level of definition. That means that if the templates were to be categorized based on the HL7 architecture, they should be placed somewhere between the domain message information model (D-MIM), which approximately corresponds to our LIM structure, and the refined message information model (R-MIM) mentioned in the HL7 version 3 standard.

We produced LIM models describing the structure of both EHRs participating in the project. During the modelling process we have realised that some of the resulting diagrams are unnecessarily complex. To solve this situation a deeper object inheritance hierarchy would be necessary. Unfortunately the HL7 version 3 standard does not cover explicitly this area (i.e. it does not explicitly allow nor restrict the inheritance of user-defined classes). As our intention was not to cause the interoperability problems by interpreting the ambiguous formulations by ourselves, we have decided to leave the schemas complex. This decision leads to a vast amount of very similar classes in the model.

Having the LIM models created we face the task of data transformation. Actual data stored in one EHR have to be transferred into another one via the HL7 message instances. The life cycle of medical data being exchanged is shown in the *Fig. 4*.

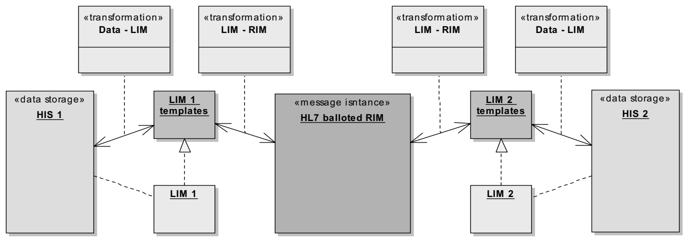


Fig. 4. Data exchange between EHRs using HL7 version 3.

The transformation process of patient data originating from a particular EHR consists of two parts. The first one deals with data transition from patient's EHR to a message that corresponds to a particular LIM template. The second part transforms the LIM template conforming the message into a HL7 message instance that uses building components from a suitable balloted RIM.

To implement the first part of the transformation process we decided to focus on a universal solution using the SOAP technology [9]. The message that corresponds to a particular LIM template is composed of logical parts – blocks. Particular blocks can be processed by a special algorithm conveniently implemented using a web-service defined in WSDL (Web Service Definition Language) file.

The security and safety requirements of the health data are an extremely important issue. The author of the information is also its authorized owner; he is responsible for its validity. The author has to keep the information according to the legislative requirements. The information about the patient is often needed by other subjects. These should keep the received information in its original form in a way allowing certifying their content and origin.

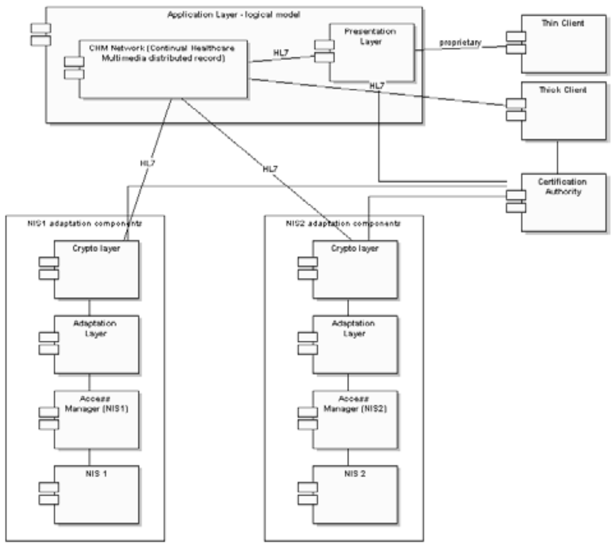


Fig. 5. The simplified logical model of the communication infrastructure.

Based on the above mentioned assumptions, the logical schema of a communication network was proposed and named the CHM network

(Continual Healthcare Multimedia distributed record). The schema has several requirements:

- The access control is maintained by the authorized owner of the information – the healthcare institution.
- The information system of the healthcare institution is equipped by the communication interface with the CHM network.
- The messages from information systems are encrypted and signed using PKI and then transferred to the recipient.
- The servers of the CHM network provide routing and transporting of the information between information systems and clients.
- All the network members use the common communication protocol – HL7 version 3.

Conclusion

The above-mentioned project aims to the development of the methods and technologies helping to ensure the continuous shared health care in information society. Based on the requirements of the communication infrastructure, clinical content, safety and security, the research and development of appropriate methodologies, models and solutions have begun. The proposed solution is based on international standards of health informatics, especially HL7, which ensures high quality and interoperability within the European Union.

Acknowledgement

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

References

[1] Zvarova J., Pribik V.: A Short Overview of Telemedicine in Czech Republic. European Telemedicine Glosary of Concepts, Technologies, Standards and Users. (Ed.: Beolchi L., Facchinetti S.) – European Communities 2003, pp. 283-285

[2] Zvarova J.: Advanced Information Technologies in the Czech Health. E-he@lth in Common Europe. (Ed.: Zielinski K., Duplaga M.) -– Krakow, Academic Computer Centre CYFRONET UST 2003, pp. 241-249

[3] Semecky J, Zvarova J.: On Regular Analysis of Medical Reports, Proceedings of Workshop on Natural Language Processing in Biomedical Applications, March 8-9, 2002, edited by R. Baud and P. Ruch, Nicosia, Cyprus 2002, pp. 13-16.

[4] Hanzlicek P.: Development of Universal Electronic Health Record in Cardiology. Health Data in the Information Society. (Ed.: Surján G., Engelbrecht R., McNair P.) – Amsterdam (Studies in Technology and Informatics), IOS Press 2002, pp. 356-360

[5] Zvarova J., Hanzlicek P., Spidlen J.: Electronic Health Record in Cardiology: Pilot Application in the Czech Republic; Achieving a

Knowledge-based Economy in Biomedicine through Informatics. – Taiwan, Taiwan Association for Medical Informatics 2002, pp. 10-13

[6] Peleska J., Zvara K., Vesely A., Zvarova J., Tomeckova M., Hanzlicek P.: Development of Electronic Form of the 1999 WHO/ISH Hypertension Guidelines. Health Data in the Information Society. (Ed.: Surján G., Engelbrecht R., McNair P.) – Amsterdam (Studies in Technology and Informatics), IOS Press 2002

[7] Zvárová J., Přečková P. (eds.): Informační technologie v péči o zdraví (in Czech), EuroMISE s.r.o., Praha, 2004, ISBN: 80-903431-1-2

[8] Tomečková M. a kol.: Minimální datový model kardiologického pacienta – výběr dat (in Czech). Cor et Vasa 2002; 44 (4) Suppl., 2002, p. 123.

[9] Seely S.: SOAP: Cross Platform Web Services Development Using XML, Prentice Hall PTR 2002, ISBN: 0130907634

Address

Ing. Petr Hanzlíček, Ph.D.
Department of Medical Informatics,
Institute of Comupter Science AS CR
Pod Vodárenskou věží 2
182 07 Prague 8
e-mail: hanzlicek@euromise.cz

Program BAYES

Heroutová Helena
Oddělení medicínské informatiky, ÚI AV ČR, Praha, Česká republika

Úvod

Bayesův vzorec se často používá v populačních etiologických studiích a v některých matematických modelech diagnostického, terapeutického či prognostického lékařského rozhodování. Bayesův vzorec

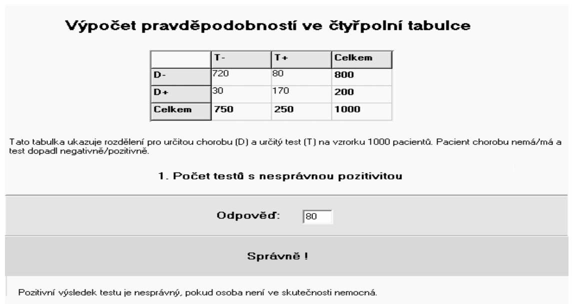
$$P(D|T) = \frac{P(T|D)\ P(D)}{P(T)}$$

tak nachází uplatnění v případech, kdy chceme vyhodnotit kvalitu diagnostického či skríningového testu. Důvodem pro provádění skríningu je, že včasná detekce nemoci vede k příznivější prognóze, protože je možno včas zahájit léčbu. U některých nemocí není skríning vhodný. Aby vhodný byl, musí být onemocnění závažné a léčba, zahájená před rozvinutím příznaků, musí být příznivější vzhledem ke snížení mortality či morbidity než v případě, že onemocnění zachytíme již v pokročilém stadiu. Navíc prevalence nemoci v preklinickém stadiu musí být v populaci, na které je skríning prováděn, dostatečně vysoká. Dalším problémem skríningu je, jak dobře vyvážit užitek včasné detekce nemoci u osob, které ji skutečně mají, a nepříznivé důsledky, které vzniknou tím, že skríningový test určí jako nemocné i ty osoby, které nemoc nemají. [6]

Z toho důvodu byl Bayesův vzorec zařazen do výuky v mnoha kurzech pořádaných EuroMISE centrem pro medicínské obory. Pro ilustraci a snadnější pochopení problematiky byl vytvořen též program BAYES.

Program BAYES

Program Bayes sestává ze dvou částí.



Obr. 1. Testování znalostí o kontingenčních tabulkách.