EXPERIMENTAL DATABASE MEDICAL SYSTEM
Data Acquisition Background and Features

Petr Maule, Jiri Polivka
Department of Computer Science and Eng., University of West Bohemia, Univerzitni 8, Plzen, Czech Republic

Jana Kleckova
Department of Computer Science and Eng., University of West Bohemia, Univerzitni 8, Plzen, Czech Republic

Keywords: Medical Software, Medical Standards, DASTA, DICOM, HL7.

Abstract: This article describes features and background of the upcoming scientific medical system. The ambitious goal of the system is to provide relevant medical data for various medical methods and experimental implementations testing. The article presents needs for such system and administrative background which it requires. The main standards which are used in medical applications are also mentioned, like the DICOM, DASTA or HL7 format with the relations to our experimental medical system.

1 OVERVIEW

Medical science is quickly evolving branch providing new treatment techniques and new diagnosis determination possibilities. Progress of software computation methods and evolution of new software algorithms plays the major role. All kinds of new techniques need to be tested and verified at large amount of suitable data. These data differs in context of methods to be verified. Some methods can strictly require one kind of input data – like one type of examination as computed tomography (CT) or x-ray images, etc. Other methods require analysis of several examination types together. Some methods can even require sequences of the same kind of examinations acquired in different time or they can even have more sophisticated needs for input data.

The goal of our ongoing medical system is to offer enough relevant data for all kind of input data requirements. If there is any new method which needs to be tested on larger amount of data, the system can provide it. It will store anonymous data coming from all our partners – mostly hospitals. The system will be sophisticated in processing of heterogeneous multimedia medical data objects. Also a knowledge-based system using the database will be proposed. Following parts describe existing examination types and storing formats, the way how the data will be transferred from our partners into the system and description of benefits for all system users.

2 MEDICAL DATA TYPES AND STANDARDS

Hospital information systems at the time of their first introduction had incompatible data formats. The same problem was with medical instruments – no mutually compatible formats. Recently the situation is much better because many standards have been established. The leading standard in digital imaging is DICOM standard which provides the same format for many imaging modalities. DICOM standard is supported in vast majority of hospital information systems and medical instruments worldwide. But there are still data which are non-imaging and need to be stored as well. Example of such data can be laboratory examinations or medical diagnosis text records, etc. For these kinds of data there are mostly used standards like HL7 and especially in the Czech Republic DASTA standard. Our system is going to support DICOM standard and DASTA standard with future plan for HL7 standard implementation.

Maule P., Polivka J. and Kleckova J. (2010).
EXPERIMENTAL DATABASE MEDICAL SYSTEM - Data Acquisition Background and Features.
In *Proceedings of the Third International Conference on Health Informatics*, pages 402-405
DOI: 10.5220/0002698004020405

2.1 DICOM Standard

The Digital Imaging and Communications in Medicine (DICOM) standard begins in 1983 when The American College of Radiology (ACR) and National Electrical Manufacturers Association (NEMA) formed a committee to develop standard unifying digital information interoperability. Current version of this standard is DICOM 3.0.

The standard covers many parts connected with medical imaging information interoperability. It defines context of information objects for different modalities as well as possible services operating with these data objects and it also defines conformance statement which must be claimed by DICOM compatible instruments. It also defines communication interface and media storing specifications.

2.2 DASTA Standard

The data standard DASTA (DS), established by the Ministry of Health of the Czech Republic, is used for transferring important data among various medical information systems in the university hospitals as well as in smaller medical facilities. The current version of the format is DS4 (DS 04.02.04) and the development is still in progress. The DS format simplifies communication between many medical information systems used in the Czech Republic, which was very complicated before (one of the medical information systems had more than 30 different protocols and diagnoses code-lists).

The first simple version of DS standard (DS 1.0) was published in the bulletin of the Ministry of Health of the CR in 1994. DS 1.0 set up the conceptual principia of the system, but it was still weak in data blocks and it didn´t solve the problematic of the laboratory segment. The next development tended to set up and improve complex communication interface with the medical laboratory information systems and to creation of the national laboratory items code-list (NLIC). The development of the first practically usable version of DS (DS1.1.) and NLIC was completed in 1997. The early version of DS standard came out from basic text files in TXT, the next versions used the XML data format (DS 2 and 3 used DTD, DS 4 are fully exploiting XML). Important versions, their launch dates and data formats are:

- DS 01.10 1. 7. 1997 "TXT"
- DS 01.20 1. 1. 2001 "TXT"
- DS 02.01 1. 5. 2002 "DTD"
- DS 03.01 2. 6. 2003 "DTD"
- DS 04.01 1. 1. 2007 "XML"

The next important moment for expansion and quality enhancement of DS standard was the supportive web-services introduction. These services offer full and actual DS4, available documentation, detailed manual and implemental programs. The regular four times a year updates are realized by the web-services too.

At the present time, DASTA standard contains not only blocks for patient data transfer, but also blocks for investigation of the drinking and supply water, diagnoses and laboratory code-lists and others. In the patient data scope, the current DS standard allows transfer of the important information from many areas. For example the patient identification data, urgent information (allergy, diagnosis), health insurance company information, anamnesis, patient medicaments, persistent and acute diagnosis, vaccinations or special data blocks. As the example of clinical events, the DS standard could be used for the laboratory work-up, RDG examination (RTG, CT and other), pulmonary function tests, consultation, ambulant and dismissory report and other.

The main advantage of the DS standard in the Czech Republic is its domestic background. On its development participated many medical software developers, so the DS meet their particular requirements. Some medical software which use the DS standard are DATAPLAN, HICOMP, ICZ, LOGIS, MEDICALC, MEDICON, PCS, SOPHIS, DS SOFT, DYNATECH and others. DS standard is used also in Slovakia, although it is not the formal standard there. The next step could be distribution of the DS to other counties in the EU.

2.3 HL7 Standard

The HL7 is another international standard in the clinical and administrative medical data domain. It was developed by Health Level Seven Inc. organization, which is accredited by ANSI (American National Standards Institute) as a SDO (Standards Developing Organization). This Health Level Seven Inc. was founded in 1987 as a not-for-profit volunteer organization like other ANSI-accredited SDOs. The standard are developed in co-operation with many various participants, who may include providers, consultants, vendors, payers and government groups who are interested in clinical and administrative standards for health care in the USA and word wide.

As the other standards, HL7 was developed like messaging protocol that enable to share and exchange sets of clinical and administrative data

among unrelated health care applications and facilities. The members of HL7 are known as the Working Group and are organized into special interest groups (SIGs) and technical committees. They are looking at end-user needs and hand it on to the attention of the HL7 Working Group. The HL7 standard is a text-based encoding system.

The first version of the HL7 standard (V.2.x) was developed and approved as an ANSI standard more than 12 years ago. The current version (V.3.0) uses an object-oriented development methodology and a Reference Information Model (RIM) to create messages. The definition of the HL7 messages was relatively free in version V.2.x whereas version V.3.0 reduces optionality of the message content which means fewer errors in the messages interpretation. In the current versions the XML protocol is used to envelope HL7 message content which increase system efficiency.

Because of its text-based format, the HL7 is still human-readable, although it is usually interpreted by machines and medical software. Several web sites provide testing HL7 messages and can be used to test a system´s ability to process the HL7 standard.

3 THE SYSTEM FILLING

As it was already said the system will be fed by our partners which will be mostly hospitals. Now at the beginning of the system implementation we have only one partner – The University Hospital in Pilsen. For all the partners we need to implement special export module to respect their information system. The University Hospital in Pilsen uses WinMedicalc (WM) information system. WM is used at many hospitals in Czech Republic. The export module can be used at all the hospitals which use WM and decide to become our partner with no need for modifications. Generally we can say that all information system mostly covers DICOM, DASTA and HL7 standards. The export module just exports data of these standards into our system with preservation of mutual relation but in anonymous form.

There are two requirements of our system on filling process. The first is that data must be able to be exported on demand and the second that all targeted data are exported automatically. There must be also duplicity removal in export module. Export module also ensures transferring from partner´s environment into our system environment which covers some kind of on-line transfers. DICOM and DASTA standard data are processed separately. The filling schema of the system is shown in the figure 1.

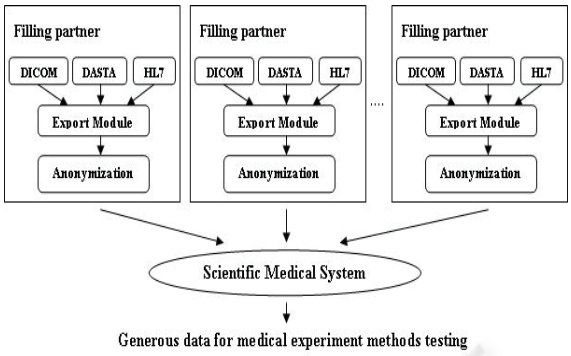

Figure 1: Filling system schema.

3.1 DASTA Exporting

The DASTA format for export is common XML format. WM provides export services which repeatedly look for new DASTA standard data in hospital information system and all found records it exports using web services or directly stores it into preselected directory at selected computer in common XML file form. At this computer our export module is running. Export module just waits for new DASTA records, process them and transfers it into our system. WM exporting ensures no duplicity of DASTA information.

3.2 DICOM Exporting

The Export module for DICOM data is more general. It can be used everywhere where DICOM standard is used. It is not depended on used hospital information system. Export module contains DICOM Application Entity (AET) which is able to communicate via DICOM standard communication interface. This module repeatedly asks DICOM image server for new records, download it, process and transfers it into our system.

Here is one feature that needs to be solved. DICOM searching for new data does not ensure unique data. DICOM communication provides query system just for data which were acquired at selected date but the data could be exported also on demand. There must be stored information in export module about already exported DICOM data to refuse all duplicities.

3.3 Patient Identification

The system preserves examinations relations. It means that we know which examination belongs to which patient in anonymizated form. During anonymization process in export module, there is

assigned unique signature for each patient. This mapping will be stored only in export module and stays inside our partner's system (in hospital). Assigned signature is randomly generated signature unique in our system. For all further data of this patient, there will be found mapping in the export module. It will use the same signature which means relations preservation.

3.4 Data Size

We are going to transfer large volume of data. Different examination means different data sizes. Probably the highest data sizes will come from DICOM file where one examination can even have about 4 GB. These data are to be transferred on-line, so there is required high-speed connection of all partners.

3.5 Drawbacks

If the patient undergo two examinations in different hospital information system scope (often means two different hospitals), it will be presented as two different patients – now we do not have enough information for mapping these like occurrences.

4 CONCLUSIONS

Our system is tending to contain large amount of different medical data provided in DICOM and DASTA standards and in the future also HL7 standard. All data will be anonymous but mutual relations will be preserved. This article shortly describes some of the examination types and storing formats which we will take into account in our scientific database medical system, like the DASTA and HL7 standard, new DICOM standard is also noticed. We introduced the approach how the data could be transferred from medical facilities and other filling partners into our system and mentioned the benefits for all system users. We will use these findings in our future research work.

REFERENCES

The DICOM Standard, http://dicom.nema.org

Vcelak P, Polivka J, Maule P, Kratochvil P and Kleckova J (2009). *Experimental database system for the vascular brain diseases research*. Frontiers in Neuroinformatics. Conference Abstract: 2nd INCF Congress of Neuroinformatics. doi: 10.3389/conf.neuro.11.2009.08.002

Data Standard MZČR, http://ciselniky.dasta.mzcr.czHealth Level 7, http://www.hl7.org