

# Privacy and Security Issues in Cerebrovascular Diseases Data Research

Petr Včelák, Jana Klečková  
*Department of Computer Science and Engineering*  
*University of West Bohemia*  
*Czech Republic*

Vladimír Rohan  
*Department of Neurology*  
*University Hospital*  
*Czech Republic*

## Abstract

*The cerebrovascular diseases research are based on real patient's data received from a collaborating centre. The goal of research is to take advantage of heterogeneous medical data relationships. Medical data relationships leads to get more data mining facilities in studying of dependencies and crucial values of parameters in a huge volume of real imaging examinations with clinical and therapeutic data. There are lot of difficulties and security issues when providing research on medical data. All real patient's identifiers must be removed for privacy reasons. A research team does not necessity know the patient identity information. We keep limited characteristics only and can be accessible to patient's medical doctors. These characteristics are useful to provide us analysis and statistics results. The aggregated results are publicly available.*

## 1. Introduction

The goal of research is to take advantage of heterogeneous medical data relationships. Medical data relationships leads to get more data mining facilities in studying of dependencies and crucial values of parameters in a huge volume of the real imaging examinations with clinical and therapeutic data. The cerebrovascular diseases research is based on real patient's data received from a collaborating centre. There are lot of difficulties and security issues to be solved in medical data research. All data used for research must be de-identified before we can use them. The experimental database concept was first announced in [17]. The knowledge base and de-identified real patient data can be used in a practical education and as a diagnostic tool.

Stroke is one of the leading causes of morbidity and mortality worldwide [7]. Large differences in incidence, prevalence and mortality have been noted between Eastern and Western Europe. This has been attributed to differences in risk factors, with higher levels of hypertension and other

risk factors resulting in more severe stroke in Eastern Europe [2]. Notable regional variations have also been found within Western Europe. Stroke is the most important cause of morbidity and long term disability in Europe, and demographic changes will result in an increase in both incidence and prevalence. It is also the second most common cause of dementia, the most frequent cause of epilepsy in the elderly, and a frequent cause of depression. [1, 12, 15]

In the European Union countries, the female and male mortality rates on cerebrovascular diseases are almost the same among persons aged 85 and over. The lowest rates are recorded in France, Slovakia, the Netherlands. The highest rates are recorded in Portugal, Greece, Latvia and the Czech Republic. The rate in the Czech Republic being 2.8 times higher than the European average. [11] The cerebrovascular diseases are the second most frequent cause of death in the Czech Republic. Cancer incidence and mortality are two to three times greater in the Czech Republic than in other developed countries in Europe. [9, 19]

## 2. Heterogeneous Medical Data

The cerebrovascular research, what we do, is based on clinical data, therapeutic data and imaging examinations. This data are not in a one same file format. Each data type has its own file format. The clinical and therapeutic data are stored in DASTA and HL7 format. DICOM format is used for imaging examinations, like CT scans and MR imaging.

### 2.1. Data Standard – DASTA

The Data Standard (DASTA) is a common widely used national electronic communication standard format of a public health service in the Czech Republic since 1994. It was developed by the Czech Public Health Informatics and Scientific Information Organisation that was supported by the Ministry of Health of the Czech Republic and ensure the medical data interoperability. Most of the medical information systems supports this format and its commonly

used for the information interchange in the Czech Republic.

The DASTA format was developed for sending laboratory results. They are structured part of the standard. The most of other items are unstructured in DASTA – clinical and therapeutic data. These are patient's identification and personal data, anamnesis, clinical events, treatment, medicaments, vaccination and urgent information, such as allergy, blood group and diagnosis. It can also contain payments, insurance company, etc. Unstructured data are a little difficult for research. [5]

### 2.1.1 DASTA Versions

It is based on the XML-based mark-up standard with XSD schema in the 4.0 version. National Laboratory Code-List is a part of the DASTA. The code-list is used for making an investigation order to a medical laboratory or getting a formalised laboratory result. The latest DASTA version is the Data Standard of the Minister of Health of the Czech Republic 04.03.01 and the National Laboratory Code-List 02.27.01. Both versions are valid since the 1st January 2010. [5]

The hospital information system Medicalc in the University Hospital in Pilsen use the previous DASTA version 03.10.01. It supports only a DTD Schema, not the XSD Schema. The National Laboratory Code-List is in 02.01.01 version.

### 2.1.2 DASTA Structure

The structure of the DASTA version 03.10.01 file format consists of a file identification such as name, date and time. Then there are DASTA and used National Laboratory Code-List versions, software name and version, destination and source department. A patient and laboratory report elements are included in the block of source department. Patient's names, birth and death date, sex, height, weight, patient assurance, urgent information, anamnesis, vaccination, medication and medical report elements are in a structured way in the patient element.

The medical report element contains plain text report and the same text in a formatted version as a Rich Text Format (RTF) document in the Base64 coding. Medical report attachments are also allowed by attached file names. CT scans in DICOM file format are attached to the medical report this way.

## 2.2. Digital Imaging and Communications in Medicine – DICOM

The Digital Imaging and Communications in Medicine (DICOM) is the industry standard for transmission of radiological images and other medical information between

computers and equipments (e.g. CT scanner and MR imaging). With DICOM format you can communicate of image data between different systems. It is applicable to a wide range of image and non-image related data exchanged in clinical and other medical environments. It becomes one of the most popular standards in medicine and is heavily used for representing and transmission radiology images scans and reporting. [10, 18]

### 2.2.1 DICOM Structure

It represents image data, so it is a binary file. Each DICOM file consists of a number of attributes, including items such as image ID, patient ID, patient name, modality, study date and time, protocol, series description, image position, slice thickness, image dimensions, the image itself and much more. In the image attribute you can store a pixel data. Mostly, each DICOM file represents single image. In fact, there can be more image frames, so it supports multi-frame images such as cine-loops. The DICOM file can contain image data in three or four dimensions. Image data size can be compressed by a variety of standards – JPEG, JPEG Lossless, JPEG2000 and Run-Length encoding (RLE). Finally, the whole data set in DICOM file can be compressed using LZW compression. [10]

## 2.3. Safe Implementation of Treatments in Stroke – SITS

The Safe Implementation of Treatments in Stroke (SITS) is an academic-driven, non-profit, international collaboration initiative by the medical professionals to accelerate clinical trials and to certify excellence in acute and secondary prevention stroke treatment. The SITS Network includes a broad range of hospitals. SITS has a base at Karolinska Institutet in Sweden. The SITS initiated an Internet-based interactive thrombosis register to serve as an instrument for clinical centres to follow their own treatment results and compare with other centres in their countries and in the collaborating countries. [16]

### 2.3.1 SITS Data Format

We consider the SITS register as important, because the clinical and treatment records are structured there. There are e.g. therapeutic details, adverse drug reaction, National Institutes of Health (NIH) data, computed tomography (CT) description, death cause description in the SITS register. This data are available only as a web page (in February 2011). It is why we have to transform web page SITS reports to our own file format. We transform data into two file formats. The comma-separated values (CSV) format and into our specific SITS-XML format with XSD Schema.

So, the data can be re-used in the cerebrovascular diseases research.

## 2.4. Health Level Seven – HL7

The Health Level Seven (HL7) is a non-profit organization involved in development of international health care standards. The HL7 organisation develops Conceptual Standards, Document Standards, Application Standards and Messaging Standards with the same name. These standards, guidelines and methodologies ease that various health care systems can communicate with each other in a uniform and consistent manner. [4]

We are interested mainly in Clinical Document Architecture (CDA) that are used to communicate documents such as physician notes and other material. The CDA is an XML-based mark-up standard intended to specify the encoding, structure and semantics of clinical documents. CDA documents must all have a header section. The header section gives details about the patient, author, date of creation etc. Always, there is a section of plain text that may optionally be divided into sections. There can be an optional section containing codes that capture essential items within the text section. [4]

### 2.4.1 HL7 CDA Structure

Now, it is in version 3 and is based on the HL7 Reference Information Model (RIM) and the HL7 Version 3 data types. This CDA document consists of a mandatory textual part and optional structured parts. The mandatory textual part ensure human interpretation of the document contents. The structured part relies on coding systems such as from SNOMED (Systematized Nomenclature of Medicine-Clinical Terms) and LOINC (Logical Observation Identifiers Names and Codes) to represent concepts and is primarily designed for software processing. [3]

The Clinical Document Architecture (CDA) is an XML document consisting of header and body part. The header part includes document type, provider, author, creation date, patient information etc. The CDA body part includes patient details, diagnosis, medications, admission details, follow-up, etc. All of this data can be present as a plain text and optionally as coded entries.

### 2.4.2 HL7 CDA Research Support

We understand DASTA disadvantages, but we have to give a preference to its processing instead of HL7, because it is used in hospital information systems in the Czech Republic now. In general, our research system is open to implement and use of mentioned HL7 standard. We bargain for the HL7 documents support. The HL7 de-identified documents

storage is supported, but data mining methods on HL7 is not implemented, at this time.

## 3. Privacy and Security Issues

Privacy and medical data security is a difficult problem. First, we have to make a boundary between privacy and the research data. We establish content of this two groups – (a) private personal data and (b) personal medical (health) data. Next, we define how to protect private personal data and some researcher's levels of medical data accessibility. Finally, we have to apply, implement and use all of this.

An emphasis is given to the privacy. That is why no medical data with patient's private data in any form can be transferred to the research database. Not only a patient privacy is important for us. It is not our research purpose to solve some medical doctor detected irregularity. What we need is the patient's medical data only for its processing. That is why we solve a similar situation with hospital staff and hospital privacy itself to the patient privacy. We have to remove hospital and hospital staff specific private data.

### 3.1. Private Personal Data

The first part of private personal data are not useful for medical research. These data or its subsets can uniquely identify the person:

- identification number (e.g. personal identification number, national insurance number, social security number),
- address (home, work, etc.),
- contact (mobile or cell phone number, e-mail, etc.),
- name and title (first name, last name, maiden name, birth name, etc.),
- insurance,
- ethnicity.

All kinds of data from the first part of private personal data must be removed without any exception.

The second part of private personal data can be a component of a medical research. Each of these data itself could not identify the person. On the other hand, you can find the person if you have all of these data available.

- age
- date of birth or death (if exist),
- gender,

- person characteristics (height, weight),
- visit date,
- ZIP code and state (only as a subset of address).

These data are useful in statistical processing and it is real to publish these data in an aggregated form. All these personal data can be visible only to the medical doctor.

For example, the known ZIP code can be used to compute patient and hospital distance with treatment delays and its dependencies on treatment success. Eventually, it can lead to a range improvements.

### 3.2. Personal Medical Data

Personal medical data are the main component of our medical research. The research can improve health for everyone and helps us learn and find out new information about cerebrovascular disease.

We can divide the medical data into two groups – (a) personal medical data and (b) hospital and other operational data like total charge, insurance or accounting price. With personal medical data or simply medical data we refer to:

- health background,
- diagnosis,
- anamnesis,
- medical report,
- imaging data (X-rays),
- laboratory result,
- medication,
- doctor or nurse note,
- urgent information,
- and vaccination.

### 3.3. Hospital and Medical Doctor Privacy

We use the same rules for medical doctors as for private personal data in . Next attribute to delete is a medical doctor specialisation. We must remove all of these data for hospital and other operational data like total charge, insurance, accounting price, etc. for a hospital privacy. Hospital and its department names are optional values that can be kept in medical data. It is fully managed by the hospital and a clinical centre which data they want to send to research database. This can protect the hospital and even clinicians.

## 4. Real Medical Data Processing and Research

### 4.1. Relationships

We profit on multiple kinds of medical data interconnected in a single database. The interconnection is build on data relationships stored in DASTA file that are illustrated on Figure 1. There the clinical and therapeutic data in DASTA and SITS-XML format you can be seen with the DICOM imaging studies. Each clinical event has its own DASTA file with the *id\_pac* unique patient identification, *note* element with the SITS treatment file identification and set of link elements to related external files. The patient identification is a clinical centre unique and specific. This patient identification has a various value for a same patient in different clinical centres because of de-identification process. The *note* element can contain one or more treatment files identification in the SITS register. The SITS identifications are comma separated. The related radiology studies in a DICOM format are linked in DASTA as *external\_file* element. Each DICOM file has own *external\_file* element in the DASTA file with file name and type. For external image data, are mostly present the thumbnails.

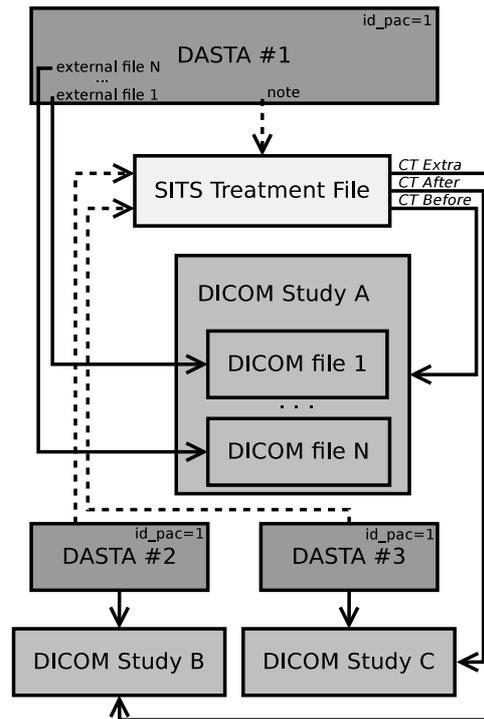


Figure 1. Medical data relationships

The Figure 1 shows that each patient’s clinical event has its own DASTA file. Only the first DASTA file refers the SITS treatment file. The SITS has three references to DICOM studies recognized by date and time equivalence of

before, after and extra CT. Not a whole DICOM study, but all DICOM files are listed in the DASTA file.

#### 4.2. Hospital and Research Team Cooperation Schema

The cooperation and medical data exchanging schema of the University of West Bohemia and University Hospital in Pilsen is on the Figure 2. The clinician records a health care data into the Hospital Information System (HIS) and the SITS register. HIS automatically exports medical data set which is de-identified before its transmission. Clinician can export patient's data set manually. Incoming medical data were pre-processed when it was received. SITS data were requested at the pre-processing time and the result is parsed and stored in research database too. Completed and uploaded data set can be processed by existing experiments and it extracts new information from the data. Medical doctor, student, developer or any other authorised research member can exploit user interface and searching for view his experiment's results.

##### 4.2.1 Feedback Enabled

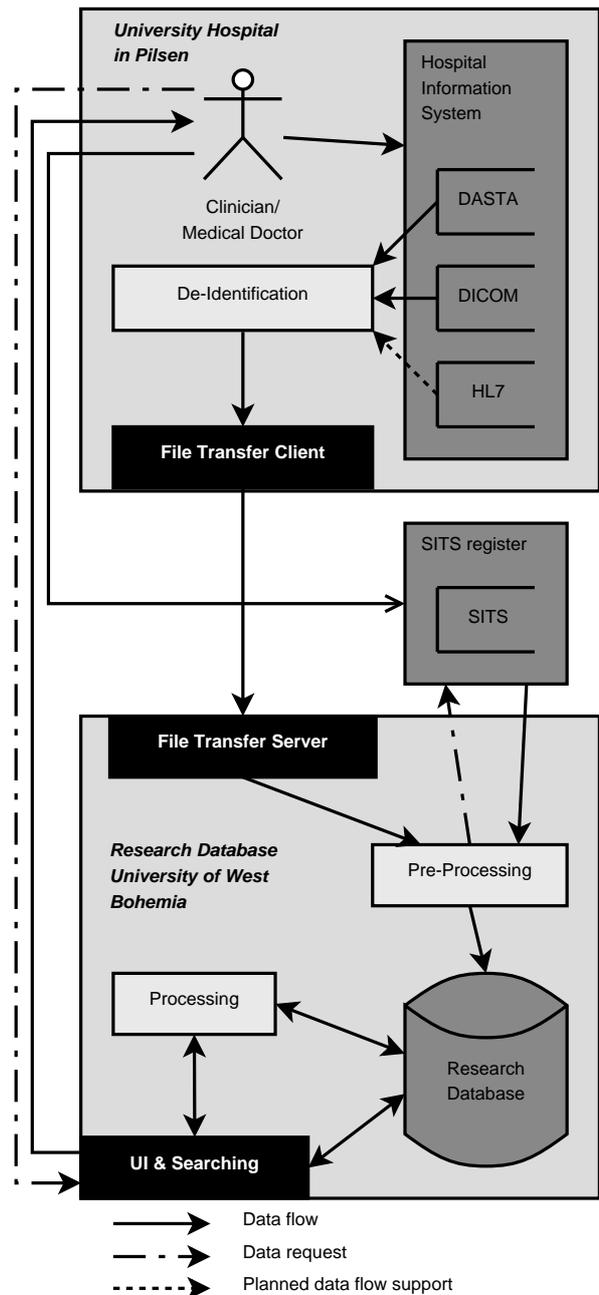
We want to have a feedback facility in research database. Some results could be useful back in hospital information system. There is no offence to privacy in a feedback support because there is no connection to the real patient in research data outside of the hospital. We keep file name and especially its export sequence number identification in research data. This export sequence number (ESN) identification is the only chance to examine the patient identity. Only, a patient's medical doctor of the right hospital can get ESN identification from the research database and then look for this identification in a hospital information system.

#### 4.3. De-Identification

All medical data are necessarily de-identified before our storage upload as you can see on Figure 2.

The medical data primarily aims a patient health care support and are stored in a hospital information system. Patient's medical data in a hospital are available to clinician working with this patient in the most of cases. If we want to use the patient's medical data for research purposes, we have to solve the problem with privacy-sensitive data. It is not important for our research to have a patient identification facility. That is why we have to automatically remove the most of these data from the clinical, treatment and imaging data.

Primarily, we definitely remove all of the private personal data as referred in 3.3. We substitute Patient's name by Anonymous name. The patient identification is substituted by a new sequence number and it's stored with private



**Figure 2. The University of West Bohemia and University Hospital in Pilsen cooperation and medical data exchanging schema**

original patient identification on the de-identification computer in hospital. This sequence identification is used for another patient's medical data. That is why we can reconstruct relationships in medical data.

Personal medical data as mentioned in 3.2 are kept in a raw clinical, treatment and imaging data. Access is driven

by registered member roles and privileges. All of these data are allowed only for a limited number of a registered and confirmed cooperating members – medical doctors. Any other registered users can view only patient's age, gender, diagnosis, medication, clinical, treatment and imaging data. Only medical doctor has allowed access to view all these patient medical data.

We use an our own de-identification software application for de-identification patient's and hospital personal data. The software application was published on [14].

#### 4.3.1 De-Identify DASTA and HL7 Formats

DASTA and HL7 formats primarily contains private personal data. We remove them as mentioned in 3.3. It is not problem to remove or change specified structured parts with this private data. There is problem only with medical report in plain text and formatted versions. In an ideal world, there would be only references to concrete person without patient personal identification, like his name – medical doctor should have to use a word patient and not his name. We live in a real world, medical doctors are a weak point in this case and we have to check for patient's name in this parts of medical report.

We remove patient private data from plain text and formatted from when we find out patient's identification, name or address. File is marked with a special warning flag when this situation occur. The formatted report text in a DASTA file format is included as an RDF document attachment in Base64 coding. It can be an RDF, DOC, PDF, TIFF or also any other file format in the HL7 format.

#### 4.3.2 De-Identify DICOM

The DICOM file format for imaging data has defined attributes with specific type and meaning. Attributes are defined by pair of group and attribute number and are well known. It is well known where a private data can occur. De-identification of DICOM file is also based on 3.3. Mostly, there are only patient's unique identification number, name, gender and date of birth.

Problem with DICOM file de-identification is with existence of (0x0028,0x0301) attribute named *BurnedInAnnotation* with value *YES*. This attribute indicates that the personal attributes are burned-in on the image. The attributes de-identification is not enough in this case and we must look for patient information at the pixel data attribute ((0x7fe0,0x0010)) to protect patient. De-identification is in an experimental version in this case. Append special warning tag similar to DASTA and HL7 format that indicates possible complication with this file. Ultrasound image mostly contains burned-in annotation.

Manufacturer's private attributes in a DICOM file can be also problematic. It is a better way to remove them too.

#### 4.3.3 De-Identify SITS

Downloaded data from this register does not contain any private data because medical doctor does not record any private data to the SITS register. We download them on the de-identification server in hospital. After the transformation to CSV and SITS-XML file format we append export sequence number identification into it. It is used to interconnect medical data later in the research database.

#### 4.4. Data Transfer

Transfer between hospital and research database must be secured because some potentially private information can exist in medical data. We use our own File Transfer Client and File Transfer Server and are on schema on Figure 2

In a real production system, we use this our client-server application [13] when we have to have a Microsoft Windows operating system on a de-identification server in a hospital. This application encrypt data on client (hospital) and sends over TCP/IP (Transmission Control Protocol/Internet Protocol) to research database for processing. Data are decrypted with a right key and stored there on a server-side. Otherwise we use stand-alone rsync tool in the real production system with on GNU/Linux de-identification server. The transfer is encrypted over SSH tunnel.

#### 4.5. Pre-processing

The medical data are checked and the relationships are extracted immediately after de-identified data upload – this is marked as a pre-processing time. At this time, we also extract all available and useful meta data about file, patient, hospital and department name, etc. The uploaded data with its meta data are stored in research database. Database structure is prepared even for missing data. Imagine the situation that DASTA file is uploaded as a first file. The database importer module read the file and recognise the SITS identification and all external DICOM files. The SITS data are uploaded too,

An unstructured clinical, therapeutic and image examinations data are indexed. An index of known medical terms and full text support is prepared. If the last file of a new data set was uploaded the importer module set the complete flag on the data set. Imported and extracted data can be used later to search proper medical data, also for evaluation or processing by experiments, when the data set is complete.

#### 4.6. Data Mining

Data evaluation and data mining are main parts of the research database. It enables to extract information from stored data sets. It could be data classification, similarity

measure methods, hypothesis, test, etc. The clinician gives instructions what to do with data and how to interpret them for research purposes – it is the description how to make an experiment. Development team implements a right algorithm for this experiment. The experiment oriented data mining can be executed repetitively. Clinician can manually start experiment execution. Automatic execution is based on a time scheduling inspired by Unix cron daemon or a new data complete trigger. The experiments supports distributed computing because of the huge volume of medical data needs to process.

The research database supports the version information of every experiment execution. Date, time, experiment, parameters, user and permission information are appended to every experiment. That is why there can be more results of the same experiment but with different e.g. parameters, user. Users have access to experiments results based on these privileges.

This data has different accessibility for research users. Data are accessible to registered researchers only. No raw data are public. Even researchers has limited access to some kinds of medical data as referred in previous sections. They have an access only for data that are necessity for their task.

#### 4.7. Results and Knowledge Base

Meta data extracted at the pre-processing time are stored together with a raw medical data and experiment meta data and results. It represents new information or knowledge of the specific experiment which can be used immediately in the next experiment. The research database knowledge base is based on modern Resource Description Framework (RDF) meta data model. It was inspired by semantically inter-operable health information systems idea that can solve many interoperability issues in the future, e.g. lack of agreed on terminologies or codes between standards (DASTA vs. HL7), inconsistent database table schema, un-cooperative health care process and different coding data structures. [6, 8]

We use the Oracle 11g database with DICOM and RDF support enabled but some data are stored in the relational database. However, we test other database, RDF storage and querying systems performance for reach the best available results. Oracle 11g database seems to be a good choice. A Virtual Private Database in the Enterprise Edition allows a fine grained access control to data – at the rows level and it is working with RDF data.

The knowledge base consists of 244 real patient's de-identified medical data sets (148 male, 96 female). The average age in this set is 68 years and it is 66.5 years for male and 71 years for female. The average male age is 66.1 and 67.2 for female. This stroke research database characteristic describes patients with stroke in years 2005–2010. This

characteristic was captured in March 2010.

For example, the research database is useful when you want to compare CT images of similar cases. Imagine, you want to select all patients with oedema and get only CT images before and after treatment of the patient with specific risk factors, grouped by similar brain area. Clinician can compare the resulting CT images without seeing a lot of non important images. It can be used for education or automatically evaluated in some experiment. The development and research can also give us better methods for working with the high data volume.

## 5. Conclusion

We describe heterogeneous medical data structure, person's privacy, security issues and relationships used for the cerebrovascular diseases research. It allows studying of crucial values and parameters dependencies in a huge clinical, therapeutic and imaging data volume. The cerebrovascular diseases research is provided in cooperation with the Department of Computer Science and Engineering at the University of West Bohemia and Department of Neurology at the University Hospital in Pilsen. The clinicians give instructions about data processing and data interpretation and computer programmers implement right algorithm. The project base is partly implemented and all ordinary heterogeneous medical data storage is supported, now. We are now working on new experiments that mine data and extend the knowledge base. The project is prepared to a new clinical centre's participants and experiment programmers. In the computer science, huge data volume tend to e.g. other storage system usage, storage performance improvements or better data index in the future.

In this paper we present data mining and knowledge base approach open to new participants and for other research topics. There are no restrictions to use research database to any other medical data research. It only depends on suitable data mining experiments.

## 6. Acknowledgements

The research presented in this paper is supported by the Students Grant Competition project number SGS-2010-077 for Applied Sciences and Computer Engineering in Biomechanics research.

## References

- [1] P. Awareness, E. Referral, P. Transfer, S. Care, D. Imaging, G. Principles, B. Tests, P. Prevention, H. Pressure, C. Smoking, et al. Guidelines for Management of Ischaemic Stroke and Transient Ischaemic Attack 2008. *Cerebrovasc Dis*, 25:457–507, 2008.

- [2] M. Brainin, N. Bornstein, G. Boysen, V. Demarin, et al. Acute neurological stroke care in Europe: results of the European Stroke Care Inventory. *European Journal of Neurology*, 7(1):5–10, 2000.
- [3] R. Dolin, L. Alschuler, S. Boyer, C. Beebe, F. Behlen, P. Biron, and A. Shabo Shvo. HL7 clinical document architecture, release 2. *Journal of the American Medical Informatics Association*, 13(1):30, 2006.
- [4] Health Level Seven, Inc. What is hl7? Access date: 2 March 2010. <http://www.hl7.org/about/index.cfm>, 2010.
- [5] Karlova univerzita v Praze – 2. lékařská fakulta v Praze (Charles University in Prague - 2nd Faculty of Medicine). Data Standard (DASTA). Access date: 2 March 2010. <http://dasta.lf2.cuni.cz/>, 2010.
- [6] R. Lenz, M. Beyer, and K. Kuhn. Semantic integration in healthcare networks. *International journal of medical informatics*, 76(2-3):201–207, 2007.
- [7] A. Lopez, C. Mathers, M. Ezzati, D. Jamison, and C. Murray. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. *The Lancet*, 367(9524):1747–1757, 2006.
- [8] D. Lopez and B. Blobel. A development framework for semantically interoperable health information systems. *International Journal of Medical Informatics*, 78(2):83–103, 2009.
- [9] Ministerstvo zdravotnictví České Republiky (Ministry of Health of the Czech Republic). Věstník č. 2/2010: Péče o pacienty s cerebrovaskulárním onemocněním v České republice. Access date: 1 March 2010. [http://legislativa.mzcr.cz/File.ashx?id=233&name=V%C4%9Bstn%C3%ADk\\_%20%C4%8D\\_02\\_2010.pdf](http://legislativa.mzcr.cz/File.ashx?id=233&name=V%C4%9Bstn%C3%ADk_%20%C4%8D_02_2010.pdf), 2010.
- [10] National Electrical Manufacturers Association (NEMA). Digital Imaging and Communications in Medicine (DICOM). Access date: 2 March 2010. <http://medical.nema.org>, 2010.
- [11] E. Niederlaender. Causes of Death in the EU. *Population and Social Conditions—Statistics in Focus 10/2006*, 2006(06):6–7, October 2006.
- [12] J. O'Brien, T. Erkinjuntti, B. Reisberg, G. Roman, T. Sawada, L. Pantoni, J. Bowler, C. Ballard, C. DeCarli, P. Gorelick, et al. Vascular cognitive impairment. *The Lancet Neurology*, 2(2):89–98, 2003.
- [13] J. K. Pavel Kodl, Petr Včelák. Medical Data Retrieval from Hospitals Over Encrypted Connection. Access date: 15 December 2010. <http://www.kiv.zcu.cz/vyzkum/software/detail.html?id=79>, 2010.
- [14] J. K. Petr Včelák. Non-Interactive Medical Data Anonymizer With DASTA and DICOM Format Support. Access date: 30 July 2010. <http://home.zcu.cz/~vcelak/sw/anonmed/>, 2010.
- [15] P. Rothwell, A. Coull, L. Silver, J. Fairhead, M. Giles, C. Lovelock, J. Redgrave, L. Bull, S. Welch, F. Cuthbertson, et al. Population-based study of event-rate, incidence, case fatality, and mortality for all acute vascular events in all arterial territories (oxford vascular study). *The Lancet*, 366(9499):1773–1783, 2005.
- [16] Safe Implementation of Thrombolysis in Stroke (SITS). About safe implementation of thrombolysis in stroke. Access date: 2 March 2010. <http://www.acutestroke.org>.
- [17] P. Včelák, J. Polívka, P. Maule, P. Kratochvíl, and J. Klečková. Experimental database system for the vascular brain diseases research. In *Neuroinformatics 2009: 2nd INCF Congress of Neuroinformatics*, Pilsen, Czech Republic, 2009. Frontiers in Neuroinformatics. Frontiers Research Foundation.
- [18] J. Walker, E. Pan, D. Johnston, J. Adler-Milstein, D. Bates, and B. Middleton. The value of health care information exchange and interoperability. *Health Affairs*, 2005.
- [19] Český statistický úřad (Czech Statistical Office). Úmrtnostní tabulky (Death-rate Statistics). Access date: 2 March 2010. [http://www.czso.cz/csu/redakce.nsf/i/umrtnostni\\_tabulky](http://www.czso.cz/csu/redakce.nsf/i/umrtnostni_tabulky), 2010.