### Information-Analytical Support to Medical Industry

Ivan Tsmots<sup>1[0000-0002-4033-8618]</sup>, Vasyl Teslyuk<sup>1[0000-0002-5974-9310]</sup>,

Anatolyi Batyuk<sup>1[0000-0001-7650-7383]</sup>, Viktor Khavalko<sup>1[0000-0002-9585-3078]</sup>

and Andreas Mladenow <sup>2[0000-0002-1233-0889]</sup>

<sup>1</sup> Lviv Polytechnic National University, S. Bandery Str., 12, Lviv 79013, Ukraine

{ivan.tsmots, vasylteslyuk, abatyuk, khavalkov}@gmail.com} <sup>1</sup> Department of Electronic Business, University of Vienna, Vienna, AUSTRIA

andreas.mladenow@univie.ac.at

**Abstract.** The list of tasks that are solved by the medical information-analytical system is determined and it is shown that it is expedient to develop such a system based on an integrated approach, covering information technologies, methods, and tools for data analytics, modeling, forecasting and decision-making. It is suggested to build a medical information-analytical system on the principles of systematicity, variable equipment composition, modularity, openness, compatibility and use of a set of basic design solutions. The component-hierarchical design method has been improved, the architecture of the medical information-analytical system has been developed, and the list of components and tasks to be solved has been defined.

**Keywords:** Medical Information Technology; Information-Analytical Systems; Component-Hierarchical Method; Neural Network; Data Mining.

### **1** Formulation of the problem

Nowadays, the effectiveness of diagnosis and treatment is largely determined by the level of information technology development. Informatization of the medical sector of Ukraine is an integral part of the state policy of the country informatization and is aimed at effective solution of health care tasks. Trends in the implementation of information technology in the medical field are associated with the creation of local information systems of diagnostic centers, telemedicine consultation systems, formation of flexible formats of medical data storage, integration of individual workplaces into a single network, first for the clinic and then for higher levels. Modern medical information technology (MIT) should be focused on the effective usage of the intellectual resource of medicine and on ensuring the availability of medical care quality to every patient regardless of residence place. Addressing these challenges requires extensive usage of telecommunications, computer and intelligence information combined with the experience of medical professionals. The effectiveness of

Copyright © 2019 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0) 2019 IDDM Workshops.

MIT depends to a large extent on the formalization of medical (diagnostic) information achieved by the collaboration of developers and physicians.

The widespread adoption of information technology in the healthcare industry requires the development of MIT components that simply adapt to the requirements of specific applications, focused on the synthesis of a wide range of medical information and information-analytical systems (IAS). The usage of pre-made components for the synthesis of IAS for specific applications will reduce the time and cost of their creation. The main task of implementing information technology in medicine is to improve the level and effectiveness of medical care. Such implementation is linked to the creation of hierarchical multilevel information systems (regional, territorial, local and individual), which should ensure effective interaction between system levels through the exchange of information in the form of information flows. Streamlining information flows is needed to improve the organizational structure of health care management, optimize health control processes, improve the workflow, and automate the process of obtaining, collecting, storing, searching, and processing data. Streamlining information flows at all hierarchical levels will increase the efficiency of the health care system and ensure the economical usage of human, financial and material resources.

With this in mind, the urgent task is to develop and use information-analytical tools to improve the effectiveness of treatment processes.

### 2 Analysis of recent research and publications

In recent years, there have been many studies and publications devoted to the issues of introduction the information technologies in the medical field, development of hardware and software components and synthesis on their basis of medical IAS [1-11]. Have a look at the most important of these publications. One of the main areas of development the medical information technologies is computerization of medical equipment, automation of collecting information process on the patient's condition, real-time processing of data and usage of the results of such processing to support decision-making. The disadvantage of existing medical information systems is the lack of intelligent means of processing and analysis of tomography data, ultrasound diagnostics, electroencephalography, electrocardiograms and radiography.

In works [1-5] the principles of construction and architecture are analyzed, the main of which is modularity of construction, openness and compatibility. Existing medical information systems do not allow the prompt storage and editing of large amounts of data and knowledge. In most cases, storing and preprocessing data comes down to the three main steps of placing, sorting, and finding need data. The development of specialized medical components of data collection and processing that provide a new level of instrumental research methods is analyzed.

The details of medical information system software development have been analyzed in publications [3-6]. It is shown that the software application is based on mathematical methods of biomedical data processing, and the development of algorithms is performed jointly with the doctors of the relevant specialties, who are task-makers. The disadvantage of the software components of existing medical information systems is that they do not fully cover the entire research process, which includes the steps of preparation, research and data processing.

From the analysis follows, that when designing modern medical IAS, it is necessary to widely use ready-made components, to create a single information space with reliable and complete information. It is also advisable to use integration as a way of organizing individual components into one system, which ensures a coherent and purposeful interaction between them, which results in high efficiency of the whole system. Integration in medical IAS is carried out in functional, informational, software-algorithmic and technical directions.

The purpose of work – choice of development principles, improvement of component-hierarchical method of designing and development of medical IAS. To achieve this goal must be performed the following tasks:

- identification of the tasks to be solved and choice of development principles of medical IAS;
- improvement of component-hierarchical design method of medical IAS;
- development of architecture of medical information-analytical system;
- determining the list of components and tasks that they solve.

## **3** Identification of the tasks to be solved and choice of development principles of medical IAS

It is proposed the development of medical IAS with an integrated approach based on new computer information technologies and covers information, organizational, technical, software and mathematical software. Information technology for the medical industry is closely linked to the accumulation and processing of information. The accuracy and reliability of diagnosis, as well as the effectiveness of treatment, depends largely on the methods of working with medical information. When working with information, the diagnosis is based on the information about the patient's running state and their interpretation. In order to accomplish such tasks, IAS components should:

- data collection, archiving and preliminary evaluation;
- generating reports on all necessary parameters of a medical card, as well as having the means to create report templates;
- automated data entry, using predictive text input and auto-fill fields;
- error detection, correctness control of entered data, monitoring of data's entry timeline, detection of unexpected or dangerous assignments, etc.
- protection of medical data from unauthorized access;
- prompt and intelligent processing of medical data, modeling;
- forecasting and planning of events of different medical nature;
- multidimensional data visualization, presentation of medical data.

In order to reduce the cost, timing and expansion of application areas, MIT is proposed to be implemented using the following principles:

- the systematic way in which links are formed between the IT components that provide integrity and interoperability with other systems;
- variable composition of equipment, which implies the presence of a kernel of information technology and variable software and hardware modules, by which the kernel adapts to the requirements of a specific application;
- modularity, which involves the development of components of medical information technology in the form of functionally completed modules that have access to a standard interface;
- openness in which medical information technologies are created taking into account the possibility of replenishing and updating functions without disturbing their functioning;
- compatibility that involves the usage of IT interfaces through which medical information technologies can interact with other systems;
- consistency of data intensities with the computing power of hardware;
- usage of a set of basic design solutions.

# 4 Component-hierarchical method of designing medical information-analytical system

The development of medical IAS is advisable to perform on the basis of the componenthierarchical method, which involves the division of the development process into hierarchical levels and types of support (algorithmic, hardware and software). To implement this approach, the decomposition method is used, which involves splitting MIT into individual components. At each level of the hierarchy, tasks of corresponding complexity are solved, characterized by both units of information and processing algorithms. The complexity of the tasks is divided into four hierarchical levels. Increasing the number of hierarchy levels corresponds to the increasing detail of algorithmic, hardware and software tools. At the higher levels of the hierarchy, information units, algorithms, software, and hardware are ordered sets of information units and compositions of algorithms, software, and hardware of the lower levels of the hierarchy (Table 1). The sequential decomposition methodology used in the development of medical information technology media reflects the top-down development process.

Hierarchical	Types of security and work in progress		
level	Algorithmic	Hardware	Software
1-st	Algorithms of MIT func-	Structure of MIT	Structure of MIT
	tioning	hardware	software
2-nd	Algorithms of subsystems	Structure of sub-	Structure of sub-
	functioning	system hardware	system software
3-rd	Algorithms for imple-	Schemes of sub-	Application librar-
	mentation of subsystem	system blocks	ies
	blocks		

Table 1. Levels and Types of development of medical IAS components

4

At the first hierarchical level of IAS development, system tasks are solved. We denote this level as  $C_{MIT}^1$ , where the unit means the first level of the hierarchy. The second level of the hierarchy consists of subsystems that provide: registration and collection of medical data; processing of medical data; data retention; analytical processing of medical data; decision-making. The third hierarchical level consists of hardware blocks and libraries, which are divided by functional features. The fourth level includes hardware and software modules that implement the basic algorithms of blocks.

The component-hierarchical structure of MIT can be described by equation:

$$C_{MTT}^{1} = \bigcup_{i=1}^{n} C_{MTT}^{2i} \bigcup_{j=1}^{m} C_{MTT}^{3j} \bigcup_{p=1}^{n} C_{MTT}^{4p}$$

where  $C_{MIT}^{2i}$ ,  $C_{MIT}^{3j}$ ,  $C_{MIT}^{4p}$  - means of the second, third and fourth hierarchical levels respectively; n – number of subsystem types; m – number of block types; h – number of module types.

The algorithms of the hardware at each level of the hierarchy are represented as functional graphs.  $F=(\Phi,\Gamma)$ , where  $\Phi=\{\Phi 1, \Phi 2, ..., \Phi n\}$  – set of functional operators,  $\Gamma$ - the law of displaying links between operators. Such representation allows us to create the list of algorithms and to determine expediency of their implementation by software or hardware.

In the treatment course of a patient by medical professionals a large amount of various information is created (text records, spreadsheets, graphics), which is processed by information technologies that combine into one integrated medical technology. Therefore, the MIT architecture under development can be regarded as an integrated technology for processing information and the means used to implement it:

 $IT_{MTT} = \{IT_{DW}, IT_{WEB}, IT_{LQC}, IT_{OLAP}, IT_{EDMS}, IT_{DM}, IT_{KDD}\},\$ where  $IT_{DW}$  – technology of information repositories (Data Warehouse);  $IT_{WEB}$  – WEB-technology;  $IT_{LQC}$  – digital signal processing (DSP) technologies;  $IT_{OLAP}$  – technology of prompt analytical processing (OLAP - On-Line Analytical Processing);  $IT_{EDMS}$  – technology of business process automation (EDMS - Enterprise Document Management System);  $IT_{DM}$  – data mining technology (DM - Data Mining);  $IT_{KDD}$  – technology that extracts new non-trivial knowledge from data in the form of models, dependencies and laws (KDD - knowledge discovery in databases).

The main role of integrated medical technology is to scientifically substantiate and find new solutions at the intersection of formal and logical approaches, taking into account the empirically descriptive nature of medicine.

### 5 Architecture and components of the medical IAS

The development of medical IAS is proposed to be carried out using a systematic approach that covers all levels of integration of processes and hardware. Synthesis of medical IAS is advisable to carry out on the basis of ready hardware and software since the development and production of new ones require considerable money and time. When choosing components, there are many factors to be considered, namely: information about the finished hardware and software components, their technical

characteristics, compliance of the interfaces with standards, opportunities for their purchase, and more. The trend of medical IAS development is focused on the transformation from a convenient tool for collecting, storing and processing data to a tool for supporting effective management decisions. For information-analytical support of treatment processes, the architecture of medical IAS was developed (Fig. 1), where DBMS - database management systems.



Fig. 1. Architecture of the medical information-analytical system

The main components of the developed medical IAS are: subsystems of registration and medical data collection; medical data processing subsystem; data storage subsystem; subsystem of analytical processing of medical data; the decision support subsystem.

**Subsystem for registration, collection and preprocessing of medical data.** The main component of medical information systems at all levels is the subsystem of collection, pre-processing and storage of medical data. Such subsystem shall provide: formalization, filtering and sorting of data; automated entry of paper documents into electronic form; registration, accounting of all volume of incoming, outgoing and internal documents; primary processing and registration of documents, data entry in databases; prompt search of document (registration number, date, authors, performers, etc.), keywords and descriptions of document fragments; optimal usage and systematization of data warehouses according to information technology needs; integration and interaction with Web-technologies, e-mail and file systems; support of various sources of information receipt; possibility to work with modern DBMS.



The structure of the subsystem for collection, pre-processing and storage of medical data is shown in Fig. 2, where ODS is an operational data source.

Fig. 2. Structure of the subsystem for the collection and storage of medical data

The main components of the medical data collection and storage subsystem are data classification tools, data structure enhancement, text parsing, data preparation, and operational data sources. It worse having a closer look at the processes that precede the recording of information in data warehouses.

*Means of medical data registration.* Much of the medical data is recorded in a variety of text documents (referrals for research, analysis, medical history, prescription, activity report), images (CT, X-rays, ultrasound) and digital signals (cardiographs).

Throughout his work, a medical professional fills in the appropriate standard forms of medical records in which. the following data are recorded:

- patient information surname, first name, patronymic, year and place of birth, nature of work, relatives;
- data on the course of treatment the results of laboratory, instrumental studies, prescribed treatment, the impact of medications, medical findings;
- data on the structure, function of medical institutions, laboratory and instrumental methods of research;
- statistical-management data (indicators of diagnosis accuracy (according to WHO classification), length of staying in hospital, degree of recovery, disagreement in diagnoses), used to calculate state medical statistics for institutions and indicators that characterize the work of a physician, department and institution as a whole;
- economic-planned data on practical and accounting activity of medical institutions.

Medical data can be obtained in the following ways:

- from analog media by recognizing the text of medical documents, digitizing images and audio data;
- directly entering data into electronic forms or from the outputs of computer tomographs, x-rays, ultrasound scanners, and cardiographs.

Obtained digital medical data is being formalized and filtered before storage. Medical data can be divided into three classes:

- highly structured (directions, prescriptions, laboratory analysis results) have a given document structure, clear formats and filling rules;
- partially structured data (some rules and formats defined but in very generalized form) results of instrumental studies and medical prescriptions;
- unstructured data (information written in arbitrary text) physicians' findings, descriptions of procedures and other information.

Means *of data structure improving*. One of the most important conditions that ensures the efficiency of medical information processing is its formalization, in which data obtained from different sources is reduced to one form. In order to identify certain trends and patterns, it is advisable to group and submit the medical data in tabular form. Obviously, the mechanisms for preparing, filtering, analyzing, and processing highly structured data are far more productive than arbitrary text analysis methods. Therefore, to effectively use medical information technology, it is necessary to increase the level of data structure.

In MIT entering of medical information is done using medical cards as well as modern medical devices (computer tomographs, fluorographs, ultrasound scanners, cardiographs, etc.), that shape the results of work electronically. The registration and collection subsystem of medical data must ensure:

8

- automated entring of paper documents into electronic form;
- registration, accounting of all volume of incoming, outgoing and internal documents;
- primary processing and registration of documents, data entry into a database;
- prompt document search according to the document attribute request (registration number, date, authors, performers, etc.), keywords and document snippets;
- optimal usage and systematization of data warehouses according to MIT needs;
- integration and interaction with Web-technologies, e-mail and file systems;
- support for various sources of information;
- possibility to work with modern DBMS.

One of the goals of MIT creating is to improve care delivery, no matter where a patient is. This is accomplished through the provision of timely and targeted highquality healthcare through telecommunications and computer technology combined with the experience of medical specialists (telemedicine).

The medical data storage subsystem is focused on the secure storage of large amounts of data, which uses data storage, that consists of two main parts: database and DBMS. Data in the subsystem can be stored in different databases and when analyzing them, there may be problems with supporting of different data formats, as well as with their encoding. This problem is solved by creating data warehouses that are domain-oriented, integrated, and immutable that support the history of the dataset.

The MIT databases contain the following: patient history; relocation of patients to a specialized unit; appointment of additional methods of medical examination, laboratory analyzes and referrals, as well as surgical intervention.

Knowledge bases used in MIT are filled with data on the experience of problemsolving by doctors in hospitals, specialized clinics, and rehabilitation centers. The experience of doctors, which is introduced into the knowledge base, is a verbal algorithm of diagnostic, therapeutic, prophylactic and rehabilitation measures with an assessment of their effectiveness.

Analytical *processing of the medical data subsystem* uses the following tools: On-Line Analytical Processing (OLAP), Data Mining (DM) and modeling.

Methods and tools for analytical data processing are based on the usage of classical statistical approaches, averaged indicators, on the basis of which to test the preformulated hypotheses and "rough" intelligence analysis. Standard static methods reject atypical observations - so-called peaks and bursts. Although such information may be interest in studying, characterizing some important phenomena. The analysis and detailed examination of such observations is useful for understanding the nature of the objects or investigated phenomena [7-11].

The *main tasks of the IAD are*: searching for functional and logical patterns in the accumulated data; finding hidden rules and patterns; building models and rules that characterize a state or predict development of certain processes. In general, the IAD process is divided into such stages [10]:

- pattern detection (free search);
- usage of identified patterns to predict unknown values (predictive modeling);
- exception analysis designed to detect and interpret anomalies in the patterns found.

Depending on the principles of working with the initial training data, all IAD methods can be divided into two large groups:

- based on the direct usage of training data;
- based on the usage of data extracted from primary data and transformed into some formal constructs.

The first group methods are used in the stages of predictive modeling and exception analysis. This group of methods includes: cluster analysis, nearest neighbor method, k-nearest neighbors method, judgment by analogy. The second group of methods includes: logical methods (genetic algorithms, decision trees, fuzzy queries and analyzes, character rules); methods of visualization; cross-tabulation methods (agents, Bayesian networks, cross-tabular visualization); equation-based methods (statistical methods and neural networks).

Modeling tools are an important means of solving many problems, in particular, modeling a functional state of a patient, his organs and systems. Using models at MIT reveals useful information that is used to make decisions. Models can be written in the form of diagrams, mathematical formulas, etc. The simplest formal description of a model is made with the following functional dependence:

$$Y = f(x_1, ..., x_n, z_1, ..., z_m, w_1, ..., w_k),$$

where Y – dependent or target variable  $x_1, ..., x_n$  – independent variables that are internal characteristics of an object of study;  $z_1, ..., z_m$  - independent variables, which are external factors that affect the object of study;  $w_1, ..., w_k$  – not taken into account characteristics or factors.

The advantage of using models to diagnose and investigate the treatment process is the simplicity of a model compared to a real object. These models allow to identify the most important factors in the object in terms of study purpose. Developed models can have different complexity, which depends on both - the complexity of a diagnosis and used treatments.

The models used in MIT can be classified according to the characteristics of the object being investigated as follows:

- dynamic and static;
- stochastic and deterministic;
- continuous and discrete;
- linear and nonlinear;
- statistical, expert, based on Data Mining methods;
- predictive, classification and descriptive.

When several different models have been developed, their selection is based on characteristics and estimates, as well as expert opinion. The main characteristics of a model that are taken into account when choosing the model are the accuracy and efficiency of the algorithm. After testing, evaluating and selecting the model, the application stage comes. At this point, the selected model is used for new data to solve tasks. The decision-making problem in information MIT has the following common features [11]:

- uniqueness of choice situation;
- cdifficult to evaluate character alternatives;
- uncertainty of the aftereffects;
- many different factors to consider when making decisions;
- the presence of a person or group of decision-makers.

In MIT, the decision-making subsystem performs the following functions:

- helps the doctor to evaluate the situation, select criteria and evaluate their relative importance;
- generates possible solutions and scenarios;
- assesses and selects solutions and scenarios;
- ensures constant exchange and coordination of information on the progress of the decision-making process;
- modeling and analysis of the possible consequences of the decisions made;
- evaluates the implementation of made decisions and conducts further training on their results.

The decisions that doctors need to make are very diverse. At one level, decisions change from "structured" to "unstructured". Structured decisions are repetitive, in accordance with the established order, decisions for which some procedure has been developed. Unstructured solutions, on the contrary, are solutions that are not in order and even unique. There are no pre-agreed procedures for making such decisions and the decision-maker must be guided by own judgments, estimations and situation. Many of the solutions are between these two groups and are called "semi-structured solutions", which means that only part of the problem can be answered clearly through a pre-agreed procedure.

At this level, when making sufficiently structured and predictable decisions, some technologies that collect, store, and facilitate access to operational-level information may be particularly useful. Operational Processing Systems (OPS) are the basic systems for operating at the operational level. They process thousands of common operations that occur daily in most organizations. In addition to the fact that these systems update records and documents in accordance with new data, they often produce convenient short reports. The information contained in such systems at the operational level is an important basis for higher-level systems.

### 6 Conclusion

1. It is advisable to carry out component development and synthesis of the IAS architecture on the basis of an integrated approach that covers information technologies, methods, and tools of analytical data processing, modeling, forecasting and decision making, and is based on the following principles of construction: systematic, variable equipment, modularity, openness, compatibility and usage of basic design solutions set.

- 2. Improved component-hierarchical method of designing the medical informationanalytical system by paralleling the design process and widespread usage of readymade hardware and software, which reduced the time and cost of designing.
- 3. To provide forecasting and search for non-obvious patterns, it is advisable to supplement traditional methods of rapid analytical processing with methods of intellectual analysis of medical data.
- 4. The architecture of a medical IAS should have a variable composition of means that assumes the presence of a kernel and replaceable modules by which the kernel adapts to the requirements of a specific medical institution.
- 5. The main components of the subsystem for collection, pre-processing and storage of medical data are: improving the structure, parsing, classification and preparation of information for recording in a data warehouse.
- 6. It is shown that increasing the efficiency of further processing of data is achieved by improving their structure, filtering and clearing errors before storing them.

#### References

- 1. Hrytsenko, V., Kotova, A., Vovk, M., Kiforenko, S., Belov, V.: Information Technology in Biology and Medicine: Lecture Course: Tutorial, Naukova Dumka, Kyiv (2007).
- 2. Korolyuk, I.: Medical informatics: Textbook, Ofort, Samara (2012).
- Medykovskiy, M., Tkachenko, R., Tsmots, I., Tsymbal, Yu., Doroshenko, A., Skorohoda, O.: Intelligent components of integrated automated control systems: monograph, Publisher of Lviv Polytechnic, Lviv (2015).
- 4. Berezsky, O., Batko, Yu., Berezska, K.: Methods, algorithms and software for biomedical image processing: monograph, Ekonomichna dumka TNEU, Ternopil, (2017).
- Dubchak, L.: Telemedicine: current state and prospects of development. In Systems of information processing, vol.1, pp.144-146 (2017).
- Rashkevych, Yu., Tkachenko, R., Tsmots, I., Peleshko, D.: Neural-Like methods, algorithms and structures for real-time signal and image processing: monograph, Publisher of Lviv Polytechnic, Lviv (2014).
- Berezsky, O., Pitsun, O., Batryn, N., Datsko, T., Berezska, K., Dubchak, L.: Modern automated microscopy systems in oncology. In: Proceedings of the 1st International Workshop on Informatics & Data-Driven Medicine, pp. 311-325 (2018).
- Berezsky, O., Verbovyy, S., Pitsun, O.: Hybrid Intelligent information technology for biomedical image processing. In: Proceedings of the IEEE International Conference Computer Science and Information Technologies (CSIT'2018), pp. 420-423 (2018).
- Tsmots, I., Teslyuk, V., Teslyuk, T., Ihnatyev, I. Basic components of neuronetworks with parallel vertical group data real-time processing. Advances in Intelligent Systems and Computing II, Advances in Intelligent Systems and Computing 689, pp.558-576, Springer (2018).
- Khavalko, V., Tsmots, I.: Image classification and recognition on the base of autoassociative neural network usage. In: Proceedings of the IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON-2019), pp. 1118-1121 (2019).
- Tkachenko, R., Izonin, I., Vitynskyi, P., Lotoshynska, N., and Pavlyuk, O.: Development of the Non-Iterative Supervised Learning Predictor Based on the Ito Decomposition and SGTM Neural-Like Structure for Managing Medical Insurance Costs. In Data, vol. 3, no. 4, pp. 1-14 (2018).

12