

## A SOLUTION FOR MEDICAL INFORMATION MANAGEMENT

RADU ADRIAN CIORA<sup>1</sup>, DANIELA GÎFU<sup>2</sup>, ADRIANA-LAVINIA CIOCA<sup>3</sup>

<sup>1</sup>"Lucian Blaga" University of Sibiu, <sup>2</sup>Romanian Academy - Iași Branch, Iași, <sup>3</sup>CMI Cioca Adriana-Lavinia, Sibiu

**Keywords:** medical data management, health information system, health profiling, big data

**Abstract:** Nowadays, the amount of data that is being generated by medical devices has exponentially increased. The aim of this paper is to develop an integrated health data management tool, that aggregates data from various sources, which are in various formats. With the aid of artificial intelligence (AI), this data will be processed and will help healthcare professionals be aware of the improvements that could make the healthcare system be more preventive, predictive and personalized. This paper introduces an integrated medical information management system – that intends to manage medical activities in hospitals, clinics and laboratories and describes its development and future directions of improvement. Furthermore, it presents a smart analysis tool that can generate both statistical data, but also infer additional information from the medical records based on natural language processing (NLP), image processing and machine learning. The novelty of the system is that it gives an overview of the patients' medical record, statistical analysis, examinations results and interpretations. Furthermore, the system is trying to predict the evolution of a disease, based on previous medical records.

### INTRODUCTION

Due to a greater interest in improving medical act, various aspects were considered. The undeniable potential of technology-based improvement was acknowledged. Furthermore, big data analytics, as an emerging discipline, is evolving into a promising direction.(1) Why? It is processing significant amounts of information from large data sets and its outcomes are generally improved compared to statistical or data mining techniques. Moreover, these benefits lead to reducing the costs of management of patient information.(2,3) Being a patient centred, the described system is based on a three-partied relation among family doctor – patient – medical institution. As Maimonides explains in (4), the doctor is faced with a patient and a medical constitution. In time the constitution and rulings change. This means that, at a given moment in time some rules apply, but others are abrogated. Keeping track of these changes, their timeframe applicability, but also patients' medical condition history can be impossible. Thus, an Integrated Medical Information Management System is more than necessary. This system apart from the tasks mentioned above can furthermore explore the big data that is aggregated from all the hugely patients' generated data set and see the big picture, which can provide aspects of evidence which otherwise would be unseeable.

Furthermore, the Healthcare Information and Management Systems Society (HIMSS) has developed the Electronic Medical Record Adoption Model (EMRAM) which is used to assess Electronic Medical Record (EMR) used in patient care, thus creating the premises for public health care institutions to become competitive, even in emerging economies, like ours.

The goal of this paper is to describe an integrated

medical information management system, together with its benefits. As such, a new technology – called AtlasMed is proposed. Its goal is to create an optimized health information system by using medical knowledge. Furthermore, it is intended to administrate medical activities in hospitals, clinics and laboratories. The novelty of the system is that it gives the possibility of gathering, aggregation, sorting of information. Moreover, stored data can be easily retrieved, search and further processed by interested parties. This is along the world-wide efforts for addressing the seemingly intractable challenge of patient safety and the growing rates of preventable deaths and injuries associated with adverse events in healthcare systems.

The paper is structured as follows: section 2 presents a short overview of existing integrated health information systems, while section 3 refers to the dataset and description of AtlasMed developed to be managed using a large medical data warehouse mentioned here. Section 4 describes the challenges that were aroused in development of the system, triggered by the different social and political changes that had an adverse impact on the management of patient information. The last section comprises of the drawing of conclusions and further directions of improvement and research.

The health care areas in Romania are still fragmented like Ellingsen and Monteiro (5) were surprised to find also in all western countries. Another paper (6) by Seoane et al., explores the biomedical data uses in drug design based on clinical records analysis and points out the value of integrating data from heterogeneous sources into a single system.

There were two major projects that we considered before analysing the subject of this paper. The first one is an open-source project called Epi Info which is a suite of statistical software tools for epidemiology that is very flexible. It is

<sup>1</sup>Corresponding author: Radu Adrian Ciora, Str. James Watt, Nr. 8, Sibiu, România, E-mail: radu.ciora@ulbsibiu.ro, Phone: +40741 381146  
Article received on 16.08.2021 and accepted for publication on 02.09.2021

developed by Centers for Disease Control and Prevention in Atlanta, Georgia, which are part of the United States federal government Department of Health and Human Services. It allows for customisable data entry, electronic survey creation and analysis. It is probably the only public domain suite that is free of charge in the field of medical records. Furthermore, it is extremely flexible, because of its three different platforms implementations: a Windows application which can be used as a stand-alone patient records database – being useful for small to medium clinics; the mobile version comes handy for remote sites, or if certain services are carried out off-site, but at the end of the day the data is dumped back into a database like the Windows one; or even better, in a cloud based solution – which was designed with large scale surveillance in mind. Its interface, apart from being customizable, can also be extended with additional information storage and retrieval tools, as database structure and interface can easily be altered from the software main graphic user interface (GUI). Furthermore, it provides an analysis module with analytic routines which include t-tests, ANOVA, nonparametric statistics, cross tabulations and stratification with estimates of odds ratios, risk ratios, and risk differences, logistic regression, survival analysis, and analysis of complex survey data. The software is an open-source project with limited support, which makes it unsustainable and unsuited for public entities.

The second project is called Integrated Health Information System (IHIS) that digitises, connects, and analyses Singapore's health ecosystem, developed by Cuspis Group. Complete IHIS product family covers all the aspects of the health information system functionalities that are used through primary, secondary and tertiary sectors, with the most relevant ones: “Complete.Primary” – for primary health care; “Complete.HIS” – for hospitals; “Complete.E-prescription” – for generating and keeping track of given prescriptions for patients; “Complete.E-Portal” – an e-portal that permits the traceability of a person's medical record; “Complete.EHR” – central fully functional electronic health record; “Complete.NHS” (“Complete.HIF”) for managing health entities inside NHS trusts or Health Insurance Funds; “Complete.Pub” for public health institutions; “Complete.E-Card” – functionality of Health ID cards.

IHIS supports most healthcare users in Singapore's health system, by means of clinical informatics, computer science, data science, mechatronics, analysis, statistical and machine learning techniques to discover insights.

Furthermore, IHIS played a key role in helping all major Singapore's public healthcare institutions become amongst the first in Asia Pacific to achieve HIMSS EMRAM Stage 6 and 7, international benchmarks for advanced technology used in patient care.

By considering the experience gained by exploring the above presented projects as well as a thorough literature review and international standards for information exchange, mandatory aspects of an integrated health information system were defined.(3,7,8) Furthermore, the challenge of fulfilling the requirements of HIMSS EMRAM international benchmarks of health entities, jump-started the process of research, specifications' definition and software specifications definition and design.

When Romania joined the EU on January 1st, 2007 a patient management system was already in place for more than a year. As a result, its medical system which at the time was mainly on paper, was at the adherence time, in the process of alignment with the EU's quality of service and practices, to enable a symbiosis between our medical records on one hand and the EU's on the other hand.

As this was not a trivial task, CNAS – the National

Health Insurance House together with Ministry of Health decided to acquire a software application that would both digitize the medical records of the people, but also to seamlessly interconnect with the EU's one. The software specifications were written by a joint effort of medical personnel, CNAS staff and IT consultants, following the national, but also the EU regulations and specifications in this field.

### Data set

We describe the dataset and the modules which comprise the AtlasMed system – used for integrating and management of medical information.

We started by compiling a corpus of medical records from on-line available repositories, with strict internal access. This data has been collected since 2006. It was stored in a traditional database, which had as backend an SQL compatible data storage solution.

The data set comprises of personal information of the patients, like name, date of birth, place of birth, home address, various identity numbers, including the social security number. It also has the medical conditions that the patient had throughout his life, until last check in the system.

For a safe patient check-in to the system, he/she must insert his own Health Card, enter the pin and then, the medical personnel are able to update his patient's medical records, both on the card, but also into the system, thus having a redundancy system what is well designed.

Furthermore, the dataset includes a very demanding storage and processing power – a medical imaging dataset (e.g. MRI, CT-scans, generated microscopic images, etc.). This dataset is not stored inside AtlasMed database, but it is interconnected with it with the aid of an API.

The best part of a medical imaging dataset is that it is annotated at acquisition time, by the medical device operator, thus leaving out the phase of data annotation. This data that is linked to the patient's personal records is timestamped and thus has well-determined spatial-temporal localization. Furthermore, cyclic imagistic data, taken at different moments in time during a patients' life span, are also stored. This enables keeping track of the patients' medical history but also allows for the traceability of different medical conditions of patients. Furthermore, with the aid of big data techniques, this vast amount of information can be used to observe various phenomena, in comparative medicine but also in seasonal and nowadays pandemic scenarios.

In the development of the system, a series of use cases were systematically analysed. Furthermore, the know-how and insights gained from studying of the existing solutions were of excellent guidance, steering the decisions in designing this medical records management system from the ground up.

*1 Baseline.* It was based on the existing frameworks presented in this paper, but also papers, best practices and reports in the field. The system's architecture and requirements were written in the contract's requirements papers. They were contracted by the implementing company of AtlasMed and then tested both in a controlled environment, but also in production, as the project stalled in the state of pilot project for a period of approximately one year. After that, with a few adjustments, it went into mass usage, being used ever since in all public sector medical institutions in Romania.

*2 Architectural Design.* The framework design is based on the agreed specifications and comprises of six modules that are loosely coupled. Interconnections among them are achieved by a message transfer system (MTS). The advantages of such an architecture are the following: each processing unit has its own private memory. Moreover, it is more efficient because of different tasks running on distinct threads or processors, thus with minimal interactions and also higher

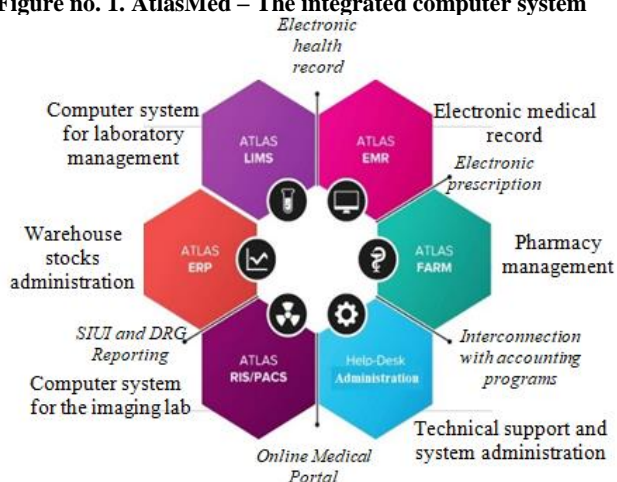
performance, due to concurrency. Furthermore, there are no memory conflicts, thus leading to low data transfer needs, which translates in less resources hungry, compared to tightly coupled counterparts.

Data acquisition is an important aspect of such a complex system. Although, the system was not designed to interface with existing systems, in the deployment stage it was necessary to import existing data from the clients. In order to achieve this a data conversion engine has been developed. This data conversion engine has some limitations, which we try to solve in this paper.

The system also has some basic data mining, as it was not specified in the initial architectural requirements. The paper comes to enhance the data analysis.

We describe AtlasMed modules (figure no. 1) spanning from directly technical integration of the information system towards integration of the medical services.

**Figure no. 1. AtlasMed – The integrated computer system**



The six components of the AtlasMed system are:

- (1) ATLAS EMR - Information system for patient administration and management of medical activities;
- (2) ATLAS LIMS - Information system for information management within laboratories;
- (3) ATLAS RIS/PACS/DICOM - Information technology management system for medical imaging laboratories;
- (4) ATLAS FARM - Information system for quantitative-value management of pharmacies;
- (5) ATLAS ERP - Information system for organizational resource management and planning;
- (6) HELP-DESK(ADMINISTRATION) - The Administration module allows configuration and maintenance of the computer system.

These loosely coupled modules are seamlessly interconnected, and their mesh covers the whole medical domain, from primary care, to laboratory, from imaging labs to pharmacies. Even though the data is in different formats and so it is not uniformly structured, the system is able to address it, store it adequately and exploit it efficiently.

**3 Data Analytics.** As we have stated in previous subchapter the system cannot automatically interface with other systems, but in the deployment stage it had to import existing records on the sites where it was implemented. This was a good opportunity for us to create an aggregated service for import/export of data between our system and other systems. In order to achieve this, an ontology was described that was able to map fields from other systems' datasets into our system's dataset, thus making the process of data transfer easier. The ontology contains list of synonym words, that are used in the medical field. Furthermore, it allows a system user to add more

terms into the ontology.

The greatest advantage of the ontology is that allows for dynamic data interrogation, enabling the user of the system to generate dynamic analytical reports, which are more context-aware than traditional database-based approaches.

Furthermore, the ontology maps the medical imaging dataset, in order to be able to browse it better and take advantage of the above-mentioned potential of ontologies.

Moreover, with the aid of computer vision techniques in conjunction with ontologies, we were able to process various data, from multiple input sources, like microscopes, MRI devices or CTs, extract relevant information and store it into the ontology's triple-store.

Further processing was achieved with the aid of random decision forests, we were able to assess if an organ – in our test case a thyroid – was healthy or not. We achieved this by submitting 121 samples labelled thyroid DICOM datasets, of which 68 were healthy samples and the rest unhealthy. The results were promising, as we were able to predict the healthiness of the tissue, with a probability of 0.78.

Healthcare today is more than ever in need of health information systems. The ever-rising expectations of governments, regulators and the public, coupled with an ageing population constantly imposes new and changing requirements to healthcare systems.

This platform is challenging, given the huge amount of data collection that is sometimes incomplete and noisy. To overcome this, certain conduct rules need to be established. Along these lines, data standardization and compulsory fields' rules definition is also a must.

As we have stated in subchapter 3.2.2 the system the system was not designed to interface with existing systems, but in the deployment stage it was necessary to import existing data from the clients. In order to achieve this, a data conversion engine has been developed. This data conversion engine has some limitations, which we try to solve in this paper.

Data mining was not specified in the initial architectural design. It was used, in order to be able to map the data between various data sources. Furthermore, it enabled for best querying capabilities, thus leading to extraction of more meaningful results on the data stored in the dataset.

In order to implement the image interpretation module, the ontology needed to be extended to incorporate image specific data. Furthermore, in order to contribute to personal medical record, by giving insides regarding imagistic information, this data needed to be interpreted. This was achieved with the aid of random decision forests.

### CONCLUSIONS

This paper presented an electronic medical records system that has ongoing challenges ahead both for software architects, but also for big data experts.

Compared to other system like Epi Info and IHIS, which are backed by classic relational databases, AtlasMed provides a more complex insight of the dataset, because of the semantic annotations and data mining techniques that are used, thus providing better insights on the data, insights which cannot be achieved in a traditional database.

Furthermore, the image processing feature offers a better view of the patient's medical record, being able to analyse medical imagistic data.

The concepts described in this paper of an integrated data management in medical care were successfully implemented in many Romanian public health institutions. With the aid of annotation, semantics, image processing and machine learning techniques there are a range of development opportunities that can be explored and from which the

Romanian health care system can benefit from.

As future opportunities of research and development, firstly, we propose a data anonymization step that is mandatory and must comply with GDPR imposed by EU's Data Protection Agency. Next step would employ data mining techniques. Thus, storing the anonymized data in a triple-store would enable a variety of opportunities for mining and querying, generating meaningful information from the large amounts of data that are stored in these databases, like epidemiological data, patient traceability or statistics.

Furthermore, the image processing module should be expanded to other organs, as well as other types of medical image data. Moreover, different machine learning techniques should be experimented with, in order to see if better results can be achieved, from the point of view of prediction accuracy, but also regarding processing time, in order to try to achieve near real-time results.

### REFERENCES

1. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. In: Health Information Science and Systems; 2014, DOI: <https://doi.org/10.1186/2047-2501-2-3>.
2. Himmel W, Rogausch A, and Kochen MM. Principles of patient management, In Oxford Textbook of Primary Medical Care; 2004.
3. Jothi N, Nur'Aini AR, Wahidah Husain W. Data Mining in Healthcare – A Review. In: Georgeta Bordea, editor. The Third Information Systems International Conference, Procedia Computer Science. 2015;72:306-313, DOI: <https://doi.org/10.1016/j.procs.2015.12.145>.
4. Bloch S. Moses Maimonides' contribution to the biopsychosocial approach in clinical medicine. Lancet. 2001;358:829-32, DOI: [https://doi.org/10.1016/S0140-6736\(01\)05974-8](https://doi.org/10.1016/S0140-6736(01)05974-8).
5. Ellingsen G, Monteiro E. The organising vision of integrated health information systems. Health Informatics Journal. 2008;14(3):223-236.
6. Seoane JA, Aguiar-Pulido V, Munteanu CR, Rivero D, Rabuñal JR, Dorado J, Paz A. Biomedical data integration in computational drug design and bioinformatics. In: Current Computer-Aided Drug Design. 2013;9(1):108-117, DOI: <https://doi.org/10.2174/15734099112089990010>.
7. Bordea G, Thiessard F, Hamon T, Mougín F. 2018. Automatic Query Selection for Acquisition and Discovery of Food-Drug Interactions: 9th International Conference of the CLEF Association, CLEF Avignon, France, September. 2018;10-14. Proceedings, DOI: [https://doi.org/10.1007/978-3-319-98932-7\\_10](https://doi.org/10.1007/978-3-319-98932-7_10).
8. Bordea G, Randriatsitohaina T, Grabar N, Mougín F, Hamon T. Query selection methods for automated corpora construction with a use case in food-drug interactions. In Proceedings of the BioNLP. 2019 Workshop, Florence, Italy, DOI: <https://doi.org/10.18653/v1/W19-5013>.