Big Data in Healthcare and Medical Applications in Romania

Adriana Alexandru, Dora Coardos
National Institute for Research and Development in Informatics (ICI Bucharest)
Bucharest, Romania
adriana_alexandru@yahoo.com)

***Abstract*—Healthcare organizations need to significantly improve the use of the health data that they hold and to share them with other organizations in order to improve the quality and efficiency of the services they provide and to scale up their learning opportunities. Big Data can unify all patient related data thus facilitating their analysis and the prediction of the outcomes. This paper introduces the Big Data concept and characteristics, healthcare data types and some major issues of Big Data and Big Data Analytics. These issues include Big Data challenges and benefits, its applications, and opportunities in medical areas and healthcare. Methods and technology progress in Big Data are also presented. Applications in Romania are briefly described**

Keywords—Big Data; Big Data Analytics; Health Care; Hadoop, HDFS, Map Reduce, Cloud Computing

I. INTRODUCTION

The science of health informatics deals with how health information is captured, transmitted and utilized for healthcare delivery [1]. Data that is rapid, complete, consistent, accurate, reliable and abundant generates useful health information.

Big Data (BD) and Cloud Computing (CC) technologies make available an impressive storage capacity to the healthcare system, providing efficiency, optimization, and care personalization. The care providers may require rapid access to computerized facilities and large storage, facilities not provided in standard IT configurations. In addition, health system data should be distributed in different geographical areas, causing significant delay in treatment and loss of time. The new technologies and storage facilities can overcome the existing deficiencies, thus facilitating those health service providers to be able to improve the services offered to their customers and patients, dispensing information more easily and improving operational efficiency at the same time.

Health care providers are now becoming more able to use data to predict health preventive interventions necessary to prescribe personalized treatments and recommend devices that provide a personalized care program. It makes a significant change from the traditional approach of dealing with medical problems, to preventive, non-invasive, driven by the use of advanced analysis based on investigations and BD.

II. BIG DATA CONCEPT

The definition of BD is debatable. According to [2], there are at least 43 definitions. Some of them are presented bellow.

Gartner defines BD as "high-volume, high-velocity, and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making" [3].

BD is the data that exceeds the processing capacity of conventional database systems. The data is too big, moves too fast, or doesn't fit the structures of conventional database architectures [4]. In 2015, it is considered that 90 per cent of the data in the world today has been created in the last two years and it is expected to grow 40 per cent per year, and 50 times by 2020 [5].

BD characteristics can be described by up to"6Vs" plus "2Vs" shared with Open Data [6-10] as described in Table 1.

BD refers to the tools, processes and procedures allowing an organization to create, manipulate, and manage very large data sets and storage facilities [11].

BD has driven the need for technological infrastructure and tools that can acquire, store, analyze and visualize large amounts of disparate structured and/or unstructured data [12].

III. BIG DATA ANALYTICS

Given the characteristics of BD, new frameworks, technologies and processes are required to manage them.

The term "*Analytics*" refers to the logic and algorithms, both deduction and inference, performed on BD to derive value, insights and knowledge from it [11]. Analytics is the process of examining and integrate vast amounts of data, from a variety of data sources and in different formats, to deliver awareness that can enable decisions in real / near real time. Various analytical methods such as data mining, natural language processing, artificial intelligence and predictive analytics can be employed to study and visualize the data. BDA approaches can be employed to recognize inherent patterns, correlations and anomalies. [12].

Table 1 – Big Data Characteristics

BIG DATA						OPEN DATA	
Volume	**Velocity**	**Variety**	**Veracity**	**Validity**	**Variability**	**Visibility**	**Value**
Data at Rest	Data in Motion	Data in Many Forms	Data in Doubt	Data Correctness	Data Changes	Data in the Open	Data of Many Values
Data come in in large scale: terabytes (TB: Approximately 10^{12} bytes), Petabytes (PB: Approximately 10^{15} bytes) and Zettabytes (ZB: Approximately 10^{21} bytes), etc of data to process	Data from continuous, time sensitive streaming flow, milliseconds to seconds to respond	Big Data extends structured, including semi-structured and unstructured data of all varieties: text, log, xml, multimedia, etc.	Quality, uncertainty due to data inconsistency and incompleteness, ambiguities, latency, deception, model approximations	Big Data need to be true and accurate.	Data flows can be consistent with regular peaks or inconsistent	Open data is generally open to anyone.	Added-value that the collected data can bring. It refers to the value that the data adds to creating knowledge.
(Records, Transaction, Tables, Files)	(Batch, Near time, Real time, Streams, Historic, Processes)	(Structured, Semi-structured, Unstructured, Multifactor, Probabilistic)	(Good, Bad, Undefined, Inconsistency, Incomplete)	(Trustworthiness, Accurate, Non-accurate)	(Consistent, Inconsistent)	(Privacy, Security, Performance)	(Statistical, Events, Correlation, Hypothetical)

Big Data Analytics (BDA) is a new approach in information management which provides a set of capabilities for revealing additional value from health information. BDA has emerged from two distinct concepts – BD and analytics. Within the health sector, it provides stakeholders (health ministries and departments, health service delivery organizations, local a regional health networks, clinicians, consumers, researchers, and vendors) with new insights that have the potential to advance personalized care, improve patient outcomes and avoid unnecessary costs [12].

BDA is unique compared to Data Warehouse (DW) /Business Intelligence (BI) systems. There are important distinctions and sufficient differentiating value between these three concepts.

Gartner defines a DW as "a storage architecture designed to hold data extracted from transaction systems, operational data stores and external sources. The warehouse then combines that data in an aggregate, summary form suitable for enterprise-wide data analysis and reporting for predefined business needs" [13].

BI is defined as "a set of methodologies, processes, architectures, and technologies that transform raw data into meaningful and useful information used to enable more effective strategic, tactical, and operational insights and decision-making" [14].

BDA solutions will not replace DW/BI, but will co-exist side-by-side to extract hidden value in the massive amount of available data.

BDA functions are unique from traditional analytic methods because they [12]: (1) Support an experimental type of analytics, whereas, traditional DW/BI and statistical analyses are based on answering known questions or hypotheses; (2) Handle open ended "how and why" type questions, whereas BI tools are designed to query specific "what and where"; (3) Process unstructured data whereas DW systems process structured, related and mostly aggregated data; (4) Process large datasets, in a distributed environment, across clusters of computers designed to scale up from single servers to thousands of machines, each having its own local computation and storage; (5) Handle the complexities of network communication, parallel programming and fault tolerance; (6) Systemically mine and flag data that is relevant for other uses and for further analytics by traditional analytics tools; (7) Leverage analytic concepts such as data mining and predictive modelling to help monitoring and forecast modelling; (8) Deliver results faster than BI systems, even in real time; and (9) Scale up to pentabytes and even zetabytes of data.

IV. USING BIG DATA IN HEALTHCARE

The revolution in healthcare data size, the diversity of data types (i.e., structured, semi structured, unstructured), the exponential increase in the speed at which it must be managed, are current problems in Healthcare Information Systems.

The percentage of programmers creating BD and analytics applications in healthcare is 4.6% in 2015 [15], as presented in Fig.1.

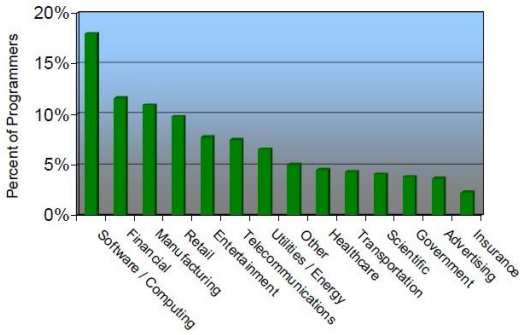

Fig.1What industries data and analytics solutions address? [15]

A. *Types of Data in Healthcare*

Specific to healthcare, the types of data available for use by BDA include [12]:

a) Genomic data – Represent significant amounts of new gene sequencing data. It refers to genotyping, gene expression and DNA sequence;

b) Streamed data – Represent significant amounts of real time data from home monitoring, telehealth, handheld and sensor-based wireless and smart devices;

c) Web and social networking-based data – Web-based data comes from Google and other search engines, Internet consumer use, social networking sites (Facebook, Twitter, Linkedin, blog, health plan websites and smartphone, etc.);

d) Health publication and clinical reference data –Text-based publications (journals articles, clinical research and medical reference material) and clinical text based reference practice guidelines and health product data;

e) Clinical data and Clinical Notes – Eighty per cent of health data is unstructured as documents, images, clinical or transcribed notes. These semi-structured (e.g., copy-paste from other structure source)) to unstructured clinical records (diagnostic testing reports, patient discharge summaries, medical images, etc.) and documents represent new data sources;

f) Business, organizational and external data – Data which previously has not been linked, such as financial, billing, scheduling, administrative, external and other non-clinical and non-health data;

g) Behavior Data and Patient Sentiment Data – Mobility sensor data or streamed data from regular medical monitoring and home monitoring, telehealth, sensor-based wireless and smart devices. Biometric data: Fingerprints, handwriting and iris scans, etc.

B. Big Data Challenges in Healthcare

1) Using complex heterogeneous patient data sources for inferring knowledge;

2) Understanding unstructured clinical or transcribed notes in the right context;

3) Efficiently handling large volumes of data from medical images and extracting potentially useful information and biomarkers;

4) Capturing new the patient data by collecting behavioral data through several sensors, as well as networking-based data related to various social interactions and communications;

5) Perform computationally intensive tasks for analyzing genomic data in combination with standard clinical data.

C. Big Data Benefits in Medical Applications and Healthcare

BD benefits in improving the quality of medical applications and healthcare can be summarized as follows [16,17]: (1) Improvement of health results through more accurate diagnosis, identification of patients at risk due to side effects and customization of care to individual patient (personalized medicine); (2) Reduction of costs through earlier detection of diseases, elimination of unnecessary care and improper claims; (3) Predicting and managing health risks, detecting healthcare fraud more quickly and efficiently; (4) Decreasing inappropriate emergency department utilization by using statistical models to identify the services or care alternatives that are more appropriated, convenient and cost efficient.

V. BIG DATA ANALYTICS IN HEALTHCARE

Effective large-scale analysis often requires the collection of heterogeneous data from multiple sources. Obtaining the 360-degrees health view of a patient benefits from integrating and analyzing the medical health records with Internet available data and readings from multiple types of meters.

A. Need for BDA in Healthcare

For healthcare, BDA represents opportunities to harness personalized care, increase the efficiency of health operations, support clinical and policy decision making, and improve patient engagement [11]. Applying advanced analytics to patient profiles can help identify the most cost effective treatments, proactively identify individuals who would benefit from preventative care or lifestyle changes. BD could help reduce waste and inefficiency in the following three areas [16]:

1) *Clinical operations*: Determine more clinically relevant and cost-effective ways to diagnose and treat patients;

2) *Research & Development*: (a) Use predictive modeling to produce a more flexible, targeted and faster R&D in drugs and devices; (b) Improve clinical trial design in order to diminish trial failures and to accelerate the introduction of new treatments on the market; and (c) Discover adverse and side effects before products reach the market;

3) *Public health*: (a) Analyze disease patterns and track disease outbreaks and transmission in order to improve public health surveillance and speed reaction; (b) Faster development of more accurate and targeted vaccines and (c) Transform very large amounts of data into actionable information.

B. Opportunities for BDA in Healthcare

The overall goals of BDA in Healthcare are [11]: (1) Take advantage of the massive amounts of data and provide right intervention to the right patient at the right time; (2) Provide personalized care to the patient; and (3) Potentially benefit all the components of a healthcare system (i.e. provider, payer, patient, and management).

Some innovative ideas and solutions to put BDA-type applications on top of the health IT infrastructure in order to derive value that might not otherwise be found are presented bellow [12]:

a) *Clinical decision support* – BDA technologies sift through large amounts of data, extract knowledge from it, and then predict results or recommend alternative treatments to clinicians and patients at the point of care;

b) *Personalized care* – Use of predictive data mining or analytic solutions for early detection of disease and diagnosis in order to highlight the most appropriated treatments;

c) *Public and population health* – BDA solutions can mine web-based and social media data in order to make disease outbreaks prediction based on consumers' search, social content and query activity. BDA solutions can also help clinicians and epidemiologists performing analyses across patient populations in order to identify disease trends;

d) *Clinical operations* – BDA can help initiatives such as wait-time management by mining large amounts of historical and unstructured data looking for patterns and modelling various scenarios to predict events that may affect wait times before they actually happen;

e) *Policy, financial and administrative* – BDA can help decision makers by integration and analysis of relevant data for key performance indicators.

VI. INFRASTRUCTURE OF BIG DATA

Traditional data processing is unable to meet the massive real-time demand of big data use. A new generation of information technology is needed to deal with BD. In Table 2, BD technologies are classified into five categories.

Table 2. Classification of BD technologies [18]

Classification of big data technologies	**Big data technologies and tools**
Infrastructure support	Cloud Computing Platform, Cloud Storage Virtualization Technology, Network Technology, Resource Monitoring Technology
Data acquisition	Data Bus, ETL Tools
Data storage	Distributed File System, Relational Database, NoSQL Technology, Integration of Relational Databases and Non-Relational Database, In-Memory Database
Data computing	Data Queries, Statistics, and Analysis, Data Mining and Prediction, Graph Analysis, BI (Business Intelligence)
Display and interaction	Graphics and Reports, Visualization Tool, Augmented Reality Technology

In this section, we will present some basic concepts of these technologies, related to data collection, storage and transferring, from a computer science.

A. Data Collection

Large scale data could be collected from many differently data sources [19]:

a) *Mobile phones* - The ubiquity of mobile phone and a large amount of data generated from embedded sensors offer new opportunities to characterize and understand user real-life behaviors (e.g., human mobility, interaction patterns). Mobile phones are also used frequently as data collection tools in the area of public health;

b) *Internet of Things (IoT)* - IoT devices generate continuous streams of data in a scalable way, and health companies must collect and handle the high volume of stream data and then perform actions on that data by using BDA;

c) *Social media* - Online social networks such as Facebook, Twitter, LinkedIn could provide a huge amount of data and large opportunities for data analytics (i.e. in disease spreading/transmission). The data from social networks could be collected by the following three ways: (1) By retrieving data shared on social network websites; (2) By asking participants about their behavior; and (3) Through deployed applications;

d) *Biomedical data from hospital and scientific community* - Data from hospital, patient data such as medical images and histories and genetic test data, large scale genomic data, etc.

B. Data Storage and Transfering

BD era is the era of NoSQL (often interpreted as NotOnlySQL) development. It is fast and cost effective. It achieves flexibility and reliability by spreading and replicating data on many different servers. NoSQL databases include MongoDB, Cassandra which is originated from Facebook, BigTable, which is developed by Google and only available to Google, and HBase, which is basically an open source version of BigTable.

Specialized BD transferring software, e.g. Aspera, is developed to maximize the transfer speed (that could reach the upper limit of an Internet connection speed) and protect the context.

C. Computational technologies

Data computing technologies cover all aspects of data processing and utilize the core techniques of BD technology. Distributed and parallel systems are needed handle large scale data.

a) Parallel computing

Nowadays, there are many different hardware architectures to support parallel computing. Two popular examples are: (1) Multicore computing (using a processor with multiple computing cores in a single chip); and (2) Cluster Computing: A cluster usually consists of set of inter-connected stand-alone computers that work together as a single integrated powerful computing resource. The individual computers with a cluster were connected by fast Local Area Networks (LAN) [16].

b) Hadoop distributed file system

With the widespread application and development of CC technologies, systems analysis based on open-source Hadoop Distributed File System (HDFS) and MapReduce data processing mode have also been widely used.

Hadoop is a software framework that can achieve distributed processing of large amounts of data in a way that is reliable, efficient, and scalable, relying on horizontal scaling to improve computing and storage capacity by adding low-cost commodity servers [18].

When dealing with the challenge of accessing BD that is stored distributed on a cluster, a common solution is HDFS, a popular type of cluster file system which is designed for reliably storing large amount of data across machines in a large scale cluster. HDFS was originally derived from Google Files System (GFS)], providing an open source cluster file system similar to GFS. The Hadoop clusters today can store PB data. It supports high aggregate data bandwidth and scale to hundreds of nodes in a cluster [16].

c) MapReduce

MapReduce is a programming model and an associated implementation for making many general data batch processing tasks, generating large data sets with a parallel,

distributed algorithm on a large-scale cluster and can have automated failover capability [18]. The MapReduce is inspired by the map and reduce functions that are commonly used in functional programming. A popular open-source implementation of MapReduce framework is Apache Hadoop.

For BD healthcare systems, the Hadoop-MapReduce framework is uniquely capable of storing a wide range of healthcare data types including electronic medical records, genomic data, financial and claims data etc. and offers high scalability, reliability and availability than traditional Database Management Systems (DBMS). In addition, intelligent functional modules such as specialized machine- learning algorithms for image analysis and recognition, diagnosis, surveillance, detection, notification etc., can be built on it [20].

d) Cloud Computing Technologies

CC and BD are complementary, forming a dialectical relationship. CC is a trend in technology development, while BD is an inevitable phenomenon of the rapid development of a modern information society. To solve BD problems, we need modern means and CC technologies. BD processing needs the support from Cloud data centers that have large-scale physical resources and CC platforms that have efficient scheduling and management functionalities [18].

BD enabled by cloud technologies could provide us new insights-clinically, operationally and in research. The concept is of storage-as-a-service CC, which provides hospitals with a BD storage capacity based on their specific demands at a low cost [21].

Besides general cloud infrastructure services (storage, compute, infrastructure/VM management), the following services are required to support BD [22]: cluster services, Hadoop related services and tools, data analytics tools (logs, events, data mining, etc.), databases/servers SQL/ NoSQL, MPP (Massively Parallel Processing) databases, registries, indexing/search, semantics, namespaces and security infrastructure (access control, policy enforcement, confidentiality, trust, availability, privacy).

VII. SOME APPLICATIONS OF BIG DATA IN MEDICAL APPLICATIONS IN ROMANIA

1) Integrated EHR System (Electronic Health Record) - SIUI

SIUI is an information system of the National Health Insurance House (CNAS), which integrates networked county health insurance houses (CJAS) and providers of health and pharmaceuticals, under the supervision of CNAS. Integration software is the unique, individualized types of providers, all CJAS unique. It has been set into full operation since 2008.

Has as Major Objectives: Management of the National Unique Health Insurance Fund (FNUASS), The uniform application throughout the country of the norms regarding the conditions for granting medical assistance within the social health insurance system in Romania; Validation according to the conditions set through the legislation of medical and pharmaceutical services paid from FNUASS; and Fraud prevention and detection.

a) Integrated EHR System (Electronic Health Record) - SIVMED

It represents a complex IT-based solution, addressed to the health units such as hospitals, groups of hospitals and policlinics. Covering the main functional aspects of the hospital environment, this software product provides solutions to both issues of medical nature and to those regarding administrative intra-hospital aspects.

Is a system formed of hardware equipments, software, people, policies and processes, which: collect data from multiple sources; is used by clinicians as a primary source of information when delivering the medical care; allows for costs evaluation with a high degree of accuracy; allows for comfortable, safe, secure and efficient access of patients and medical staff to data and information of interest.

It is based on modern technologies, offering immediate and long-term benefits, as well as the automatization of repetitive processes, increased workflow efficiency, security and interoperability, including interoperability with health insurance systems.

2) Electronic Prescription Information System - SIPE

It aims to provide one of the essential components of the Informatics Platform Health Insurance in Romania, as a step towards alignment with the trends in the European Union.

Is a natural extension of the current CNAS SIUI system providing all functionalities necessary for the introduction of electronic prescribing compensated and free recipes, borne partly or wholly from FNUASS.

It became operational on 01/07/2012.

3) Armonia®sanita

Is a hospital information system successfully implemented by some of the biggest medical service providers in Romania. It is available in Cloud.

With modules covering most work flows in a hospital, from patient scheduling, hospitalization and laboratory, pharmacy and even surgery management, ARMONIA®sanita contributes to higher efficiency and increased quality in healthcare services.

With the use of ARMONIA®sanita, the patients' waiting time for tests, investigations and consultations is reduced, the quality of healthcare services to western standards is increased and the complex reporting, specific to healthcare service providers is facilitated.

The electronic prescription is integrated with the medical, financial and administrative modules of the information system for hospitals. The direct communication of the electronic prescription with SIUI-SIPE, allows for the direct export of series and numbers for electronic prescriptions and pre-typed electronic prescriptions.

4) Integrated Emergency Service System

Coordinate emergency services at national level and enhance response actions of operational units.

Manage 1,400,000 victims each month.

Owing to the implemented technologies, manage to reduce the time for victim spotting on the map and we increased the response capacity of search teams within the entire territory of Romania.

5) National Registry of Patients with Rheumatoid Arthritis

Association of Romanian Register of Rheumatic Diseases and Ymens Cloud Company developed and implemented a cloud solution of reference for the public health sector in Romania. Developed nationwide, the system is dedicated to the automation of the flow of the dossiers of patients with rheumatoid arthritis

The main objective was to deploy a SaaS cloud to ensure a fast flow of files and allow patients access to medication in an accurate and controlled through online processes, approval or rejection therapies. The direct users are generally physicians, and they needed an intuitive, easy to use system, allowing the easy and accurate collection of data on the evolution of specific diseases in Romania and immediate analysis reports on treatment outcomes, as well as operation services, maintenance and technical support;.

It received Romania EuroCloud Award 2014.

6) StarVault e-Health

The platform that allows doctors and medical centers to interact with patients in a safe and friendly environment in order to make permanent exchange of medical information from medical systems, but also from patients and doctors input.

Provides access to content for archiving data, in order to: Deliver advanced storage and data management capabilities; Provide protection, conservation and lifecycle management for compliance requirements, internal controls, records management and legal regulations; Provide advanced search features, Supports rapid growth of unstructured data; Ensures scalability for cost control; Support thousands of customers in a single system; and Improve operational efficiency.

EuroCloud Award 2015 Received Romania.

VIII. CONCLUSION

BD is based on data obtained from the whole process of diagnosis and treatment. BD has great potential to improve medicine, and to guide clinicians in delivering value-based care. BD offers a lot of benefits such as: disease prevention reduced medical errors, right care at the right time and better medical outcomes. Challenges in medical applications and healthcare include: aggregating and analyzing unstructured data, indexing and processing continuously streaming data, privacy, and information.

BDA presents new opportunities to combine health science domain expertise with data science to provide the insights that to meet the decision makers needs to transform health delivery. BDA is moving forward quickly in population health and quality measurement.

With the help of Hadoop, the effective citizen care management can be achieved by providing an effective data driven services to citizens, and by predicting their needs

REFERENCES

[1] American Health Information Management Association, "What is Health Information?", http://www.ahima.org/careers/healthinfo

[2] J. Dutcher, "What Is Big Data?", Berkeley School of Informatics, Sept. 2014.

[3] Gartner, "IT glossary: big data" [webpage on the Internet], Stamford, CT:; 2012, Available from: http://www.gartner.com/it-glossary/big-data.

[4] E. Dumbill, "Making sense of Big Data. Big Data", 2013.

[5] M. de Waal-Montgomery, "World's data volume to grow 40% per year & 50 times by 2020",2015,https://e27.co/worlds-data-volume-to-grow-40-per-year-50-times-by-2020-aureus-20150115-2/.

[6] P. Russom, "Big data analytics", TDWI Best Practices Report, Fourth Quarter, 2011.

[7] C. Eaton, D. Deroos, T. Deutsch, G. Lapis and P. Zikopoulos, "Understanding big data", McGraw-Hill Companies, 2012.

[8] O'Reilly Radar Team, "Planning for big data", O'Reilly, 2012.

[9] P. Zikopoulos, C. Eaton, D. de Roos, T. Deutsch, G. Lapis, "Understanding big data: Analytics for enterprise class hadoop and streaming data", McGraw-Hill, New York, 2012.

[10] Forbes, "What is big data?", http://www.forbes.com/sites /lisaarthur/ 2013/08/15/what-is-big-data/.

[11] J. Sun, Reddy C.K., "Big Data Analitics for Healthcare", Tutorial presentation at the SIAM International Conference on Data Mining, Austin, TX, 2013, https://www.siam.org/meetings/sdm13 /sun.pdf.

[12] Canada Inforoute , "Big Data Analytics in health", White Paper, Full Report, April 2013.

[13] Gartner Research, "Magic quadrant for business intelligence, Information governance - 12 things to go in 2012", Gartner_Keynote_Big_Data_ and_Master Data Management.

[14] Frost & Sullivan White Paper, "Drowning in Big Data? Reducing Information Technology Complexities and Costs For Healthcare Organizations", 2012. http://www.emc.com/collateral/analyst-reports/frost-sullivan-reducing-information-technology-complexities-ar.pdf

[15] L. Columbus, "2015 Big Data Market Update", Forbes/Tech, 2015, http://www.forbes.com/sites/louiscolumbus/2015/10/11/2015-big-data-market -update/#bde9d0051704.

[16] L. Wang, C. A. Alexander, "Big Data in Medical Applications and Health Care", American Medical Journal, Vol. 6, Isuue 1, pp. 1-8, DOI: 10.3844/amjsp, 2015.

[17] W. Raghupathi, and V. Raghupathi, "Big data analytics in healthcare: Promise and potential", Health Inform. Science Syst., vol. 2, pp. 1-10, DOI: 10.1186/2047-2501-2-3, 2014.

[18] W. Tian, Y. Zhao, "Optimized Cloud Resource Management and Scheduling. Theory and Practice", Elsevier Inc., 2015.

[19] T. Huang, L. Lan, X. Fang, P. An, J. Min, F. Wang, "Promises and Challenges of Big Data Computing in Health Sciences", Big Data Research, vol. 2, pp. 2–11, 2015.

[20] C. Ngufor, J. Wojtusiak, "Learning from largescale distributed health data: An approximate logistic regression approach", Proceedings of the 30th International Conference on Machine Learning, Atlanta, Georgia, USA, JMLR: W&CP, pp: 1-8, 2013.

[21] J.C. Hsieh, A.H. Li, C.C. Yang, "Mobile, cloud and big data computing: Contributions, challenges and new directions in telecardiology", Int. J. Environ. Res. Public Health, vol. 10, pp. 6131-6153, 2013, DOI: 10.3390/ijerph10116131

[22] M. Turk, "A chart of the big data ecosystem", 2012.