# Healthcare, Big Data and Cloud Computing

ADRIANA ALEXANDRU[1], CRISTINA ADRIANA ALEXANDRU[2], DORA COARDOS[1], ELEONORA TUDORA[1]
[1] National Institute for Research and Development in Informatics, 8-10 Averescu Avenue, Bucharest, 011455 ROMANIA
[2] School of Informatics of the University of Edinburgh, Informatics Forum, 10 Crichton Street, Edinburgh EH8 9AB, UK
adriana_alexandru@yahoo.com

*Abstract:* - The rapidly expanding field of *Big Data* is a result of the data deluge generated by the rapid increase in the use of mobile devices and social media. It has provided tools to collect, store, manage, and analyze huge volumes of data (either structured, semi-structured or unstructured) produced by current healthcare systems. *Big Data Analytics* has started to play a crucial role in the evolution of healthcare practices and research, being now applied towards aiding the process of care delivery and disease exploration. *Cloud computing* enables all Big Data operations through the provision of large storage and processing power. This paper introduces all of these notions in the context of healthcare. It discusses the advantages, but also challenges brought by Big Data to this field. It also proposes a Big Data Architecture for healthcare. Finally, it presents some of the areas of application of Big Data to healthcare in Romania.

*Key-Words:* - Big Data, Big Data Analytics, Cloud Computing, Healthcare, semi-structured data, unstructured data, NoSQL databases, Hadoop

## 1 Introduction

Healthcare aims to maintain and improve human health via diagnosis, treatment and disease prevention. As motivated by patient care, auditing, administration and requirements to conform to standards and regulations, it generates massive amounts of data - Big Data (BD) - every day [1]. BD has the potential to support many medical and healthcare operations, including clinical decision support, disease surveillance and population health management [1] [2].

Big Data Analytics (BDA) shifts data management and its techniques from structured to semi-structured or unstructured data and from a static terminal environment to a ubiquitous cloud-based environment. BDA applications in healthcare use the rapid increase in data volumes to discover associations, patterns and trends in the inherent complexity of the data, and to find actionable insights for smarter decision making [1]. They thus have the potential to improve the quality and reduce the spiraling costs of healthcare delivery. Healthcare organizations must invest in the right tools, infrastructure and approaches to be able to exploit this potential [1] [3].

Cloud computing is the most appropriate architecture for large-scale storage and complex processing as required for BD and BDA. Its advantages include flexibility, security, parallel processing, scalability and resources virtualization [4]. Cloud computing can reduce the costs of automation, informatisation and infrastructure maintenance, meanwhile improving operational efficiency and user access [4] [5].

The push for BDA in healthcare and the increasing importance of cloud services have provided new perspectives to healthcare [6]. More and more medical data is uploaded to the cloud and shared by healthcare professionals, and more and more analytics is performed on BD medical repositories.

This paper provides an overview of the field of BD and its applications in healthcare. Section 2 offers a more detailed introduction to BD, BDA and cloud computing, and the opportunities that they bring in the context of healthcare. Challenges also exist, and they are discussed in Section 3. In Section 4, we propose a BD architecture for healthcare. Finally, Section 5 presents some of the applications of BD in healthcare in Romania.

# 2 Background

## 2.1 Big Data Definitions and Characteristics in Healthcare

Since being coined in 1997 by NASA researchers, the term BD has been redefined repeatedly during the years [7]. The following are some of its most popular definitions:

- "Data of a very large size, typically to the extent that its manipulation and management present significant logistical challenges; also, the branch of computing involving such data" (Oxford English Dictionary) [8]
- "Datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze" (Manyika et al.) [9]
- "Things one can do at a large scale that cannot be done at a smaller one, to extract new insights or create new forms of value" (Mayer-Schönberger and Cukier) [10]

In healthcare, BD includes patient-related data from electronic health records (EHRs), computerized physician order entry systems (CPOE), clinical decision support systems, medical devices and sensors, etc., and less patient-related data such as emergency care, news feeds and research articles [1]. This data is tremendous due to its volume, diversity of data types and speed at which it must be analyzed, which makes it difficult (or even impossible) for traditional software and hardware to handle it [1] [11].

The characteristics of BD can be synthesized by 7 Vs ([12]), which are summarized in Figure 1:

1. ***Volume***: the considerable increase in data volume caused by both traditional, as well as new, date types becomes a storage problem, but also affects the complexity of data analysis [13]. The volume of worldwide healthcare data is estimated to grow to 25,000 petabytes (1015 bytes) by 2020 [2]. This is caused by the conversion of existing data from paper into digital form, but also by emerging new forms of data (e.g. 3D images, genomics, etc.) [1][2]. Cloud-based platforms for the more effective collection, storage and management of these large volumes of data are needed.
2. ***Variety***: nowadays, new channels and emerging technologies (e.g. social media, the Internet of things) generate more and more diverse semi-structured or unstructured data [13].
   In healthcare, structured and semi-structured data is continuously generated through the movement of information from paper medical records to EHRs and electronic medical records (EMRs) [2]. It also includes readings from medical devices, some of the clinical data, electronic accounting and bills, statistics used to calculate risks. Unstructured data can be found in paper medical records and prescriptions, handwritten notes, X-rays and other images.

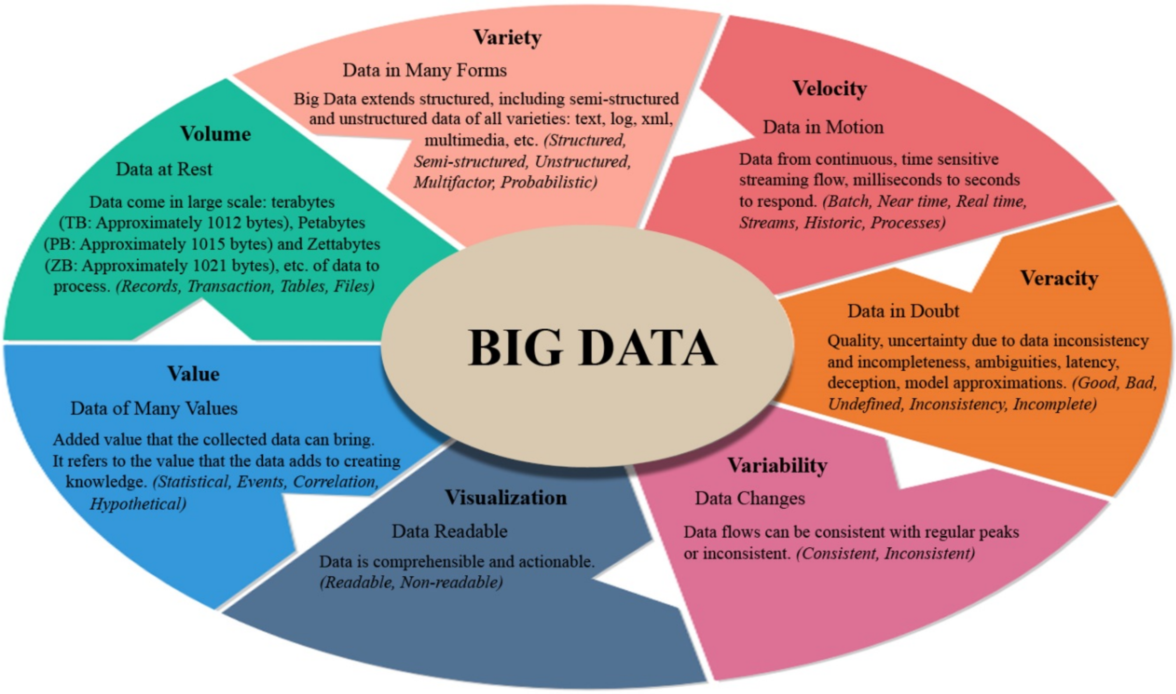


Fig. 1 Big Data Characteristics

Moreover, new forms of data are now being generated by fitness devices, genetics, genomics, social media, and through research. Analyzing and processing this large volume of very variable data in order to extract new insights becomes a great challenge [14].

3. ***Velocity***: refers to the speed with which data is generated, retrieved, processed and analyzed. This speed is continuously increasing due to real- time generation algorithms, requests combining data flows with business processes and decision making processes [1].
   In healthcare, data can be static (e.g. paper medical records, X-ray images, etc.), medium-velocity (in case of repeated measurements, EKGs, etc.), or real-time (e.g. bedside and operating room monitors) [2]. The generation of up-to-date results from real-time or near real-time processing is important, for example, in clinical decision support, for being able to make the right decisions rapidly [14].
4. ***Veracity***: refers to the accuracy and compliance of the data. This is particularly difficult for BD, because data may come from various sources which cannot guarantee quality and compliance to a particular presentation format [13].
   Data veracity is a particular concern in healthcare because it affects patient safety, but also because of the risk of low quality, or even incorrect data (often in the case of unstructured data, e.g. misinterpreted handwritten prescriptions) [1]. High-quality data is a prerequisite for improving the quality of care (by making more accurate diagnostics and better choices of treatments), while reducing costs.
5. ***Variability***: refers to how interpretable the data is. For example, sophisticated algorithms are necessary in order to deduct the correct meaning of social media comments.
6. ***Visualization***: refers to the readability and accessibility of the data presentation [13], which requires numerous spatial and temporal parameters and relationships between them.
7. ***Value***: refers to the possibility of creating new knowledge and economic value by exploiting the data.

Tools and skills are needed to efficiently analyze the large amount of healthcare data in order to extract its value and thus improve health outcomes.

## 2.2 Big Data Analytics in Healthcare

The following are two of the definitions which were provided for BDA in healthcare:

- "the systematic use of data and related clinical and business (C&B) insights developed through applied analytical disciplines such as statistical, contextual, quantitative, predictive, and cognitive spectrums to drive fact-based decision making for planning, management, measurement and learning" (HIMSS) [15];
- "the use of data, information technology, statistical analysis, quantitative methods, and mathematical computer-based models to help health care providers gain improved insight about these patients and make better, fact-based decisions" (Raghupathi and Raghupathi) [16].

*Applications for BDA in healthcare* aim to improve patient safety and clinical outcomes, while also promoting wellness and disease management. [17].

***A four stage model of BDA in healthcare*** consists of [16]:

1. *Descriptive analytics*: examines data to understand past and present healthcare decisions and make new informed ones. Its models can categorize, define, combine and classify the data to extract useful information. Charts and reports are often used for data visualization.
2. *Predictive analytics*: examines old or summarized healthcare data in order to identify patterns of relationships which it can extrapolate to predict the future [16]. It enables understanding of why it happened and what is likely to happen in the future. It can help improve patient care, the management of long term conditions and healthcare administration [18].
   Data mining can be used to identify hidden patterns in large volumes of health data, and thus anticipate medical risks, predict outcomes for patients and improve health related services [16].
   Socioeconomic data can be useful for predictive analytics [19]. For example, it can help organizations to predict missed appointments for patients living in remote areas with a certain postal code where owning a car is unlikely, and prevent them (thus reducing costs) by sending a taxi to pick such patients up for their appointments. This also helps reduce the number of hospital readmissions.
3. *Prescriptive analytics*: uses data or information, but also health and medical knowledge to solve problems which involve too many alternatives, making descriptive or predictive analytics infeasible [16]. It is often used in healthcare, for example to help decide amongst different

treatments or interventions, but also in personalized and evidence-based medicine.

4. *Discovery analytics*: uses knowledge about knowledge to identify previously unknown facts from data and improve the future [16]. It can help, for example, discover new diseases or medical conditions, drugs and treatments.

Many applications use all of the four types of data analytics, albeit by using different models and tools [16].

### 2.3 Interdependency of Cloud Computing with Big Data Technologies

Cloud computing is defined by Sultan as “an area of computer science in which highly scalable IT capacity is provided in the form of services delivered via the Internet to numerous external users” [20]. It is a fast-growing technology and, according to NIST, “a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction” [21].

Cloud computing is an enabler of all BD operations [22]. In fact, the two terms are so interrelated that BD experts often consider them as indistinguishable.

Cloud computing offers the following important advantages for healthcare [23]:

- It enables access to large storage and processing power on demand;
- It allows the movement of EHR, genomic and radiology data through BD sets;
- It facilitates the sharing and access to EHRs for care providers from different geographic locations, thus reducing the need for patients to be re-assessed;
- It allows for BDA to be performed in order to obtain actionable insights.

**Service Models**

The cloud model is composed of three service models: Software as a service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS) [22]:

- *Software as a Service (SaaS)*: the provider manages and controls a cloud infrastructure, and also runs applications on it. The consumer can use these applications by means of a client interface, and has limited access to some of their configuration settings.
- *Platform as a Service (PaaS)*: consumer-developed or purchased applications are run on the cloud infrastructure with support from the provider. The consumer has control over these applications, and sometimes over external setups. The provider is the one to manage and control the cloud infrastructure.
- *Infrastructure as a Service (IaaS)*: the consumer provides computing resources (e.g. processing, storage) and can run software onto the cloud. While the cloud infrastructure is still managed and controlled by the provider, the consumer has control over the software’s environment, sometimes including network components.

**Deployment Models**

Depending on the availability of the cloud infrastructure, the following deployment models can be defined: [22]:

- *Private cloud*: the cloud infrastructure is intended for a single organization, and may be managed by it or by a third party.
- *Community cloud*: the cloud infrastructure is used by several organizations with common concerns. They may also manage it, or allocate this task to a third party.
- *Public cloud*: an organization provides the cloud infrastructure as a service for open use.
- *Hybrid cloud*: the cloud infrastructure consists of several separate clouds of the above types, which are linked by means of standardized or proprietary technology.

In healthcare, a certain deployment model is chosen based on the type of application and its requirements [23]. Private or hybrid clouds are usually used for clinical applications such as EHRs and physician order entry systems because of the criticality of privacy, security and availability, while public clouds may be used for nonclinical applications (e.g. automatic patient billing).

## 3 Big Data Challenges in Healthcare

The rapid advances in the field of BD have led to important opportunities, but also to a series of challenges, some of which have important implications for healthcare:

### 3.1 Interpretation and Correlations

The increase in the amount of data and the speed with which it is generated by orders of magnitude has several negative implications [24]:

- A reduced need for sampling, and thus discordance with the majority of statistical approaches developed in an era of “sample sizes” and necessity to find new or improved data analysis methods;
- The de-prioritization of very well curated data for obtaining higher sample and effect sizes [24] [10];
- Possibility to incorrectly discover patterns because of the large number of emerging connections [24] [25];
- Possibility to fraudulently produce a significant effect by using the increased amount of data to identify those combinations bringing marginal changes which can be multiplied given the amount of instances or running time [24] [26];
- Over reliance on correlations due to the lack of causality or strong theories [24] [27].

### 3.2 Standards and Interoperability
BD increases the need for interoperability, and thus standardization, in healthcare [24] [28]. However, standardization is challenging due to data fragmentation, duplication, format incompatibility between systems, and the fact that some organizations still use their own data and information infrastructures.

### 3.3 Privacy and Security
More and more people have access to, and use, the increasing amount of healthcare data, which demands assurances as to privacy and security in light of existing laws [24].

To ensure security, data access on the cloud should be controlled by group, role and function for the different stakeholders [14]. Moreover, security must be ensured even once the data leaves the cloud, which is where BD can also help.

### 3.4 Data Expertise and Infrastructure
Silos are very used in healthcare, but rely on very different content in different organizations (e.g. paper or electronic charts) [29]. Expertise is required to harmonize such content and build data warehouses in order to facilitate their analytical use [25]. However, most healthcare organizations lack such expertise and an understanding of the advantages of BD ([29]), or the funding necessary to develop expertise in innovative data management solutions ([24]).

### 3.5 Timeliness
Time is of the essence in obtaining data for clinical decisions [14]. BD speeds up decision support, and may make it more accurate due to the reliance on huge amounts of data. However, attention to the data and to queries are very important in order to ensure that time constraints are respected, while still getting accurate answers.

### 3.6 The IT leadership
To make the most of BD, the IT leadership should support medical experts in sharing findings and acting on them to improve patient care [29]. Medical experts should exchange knowledge from different clinical clouds and use a special cloud service which can be integrated into clinical workflows.

## 4 A Big Data Architecture for Healthcare
Figure 2 presents a BD architecture for healthcare which was inspired from that by Wang et al. ([30]) and previous work ([31]). Its technical side consists of four important layers [30]:

1. ***The data layer***: includes hospital or external data sources for structured data (e.g. EHR records), semi-structured data (e.g. health monitoring device logs) or unstructured data (e.g. clinical images);
2. ***The data aggregation layer***: performs three important operations on the data from the various sources [30]:
   - *Data acquisition*: data is collected from various communication channels, with different frequencies, sizes, and formats (texts, images, videos), by using sensors connected to a wireless network or mobile devices;
   - *Data transformation*: ETL (Extract, Transform, Load) tools enable data pre-processing into suitable formats and types. For example, structured data could be extracted from an EHR system, its format could be changed, it could be sorted (e.g. by patient name) and then validated to respect quality requirements.
   - *Data storage*: depending on the content, the data is loaded into appropriate databases for further processing and analysis [30]. This can be performed in batch processes, in near real time or in real time [13]. With the growth in data volume, traditional relational

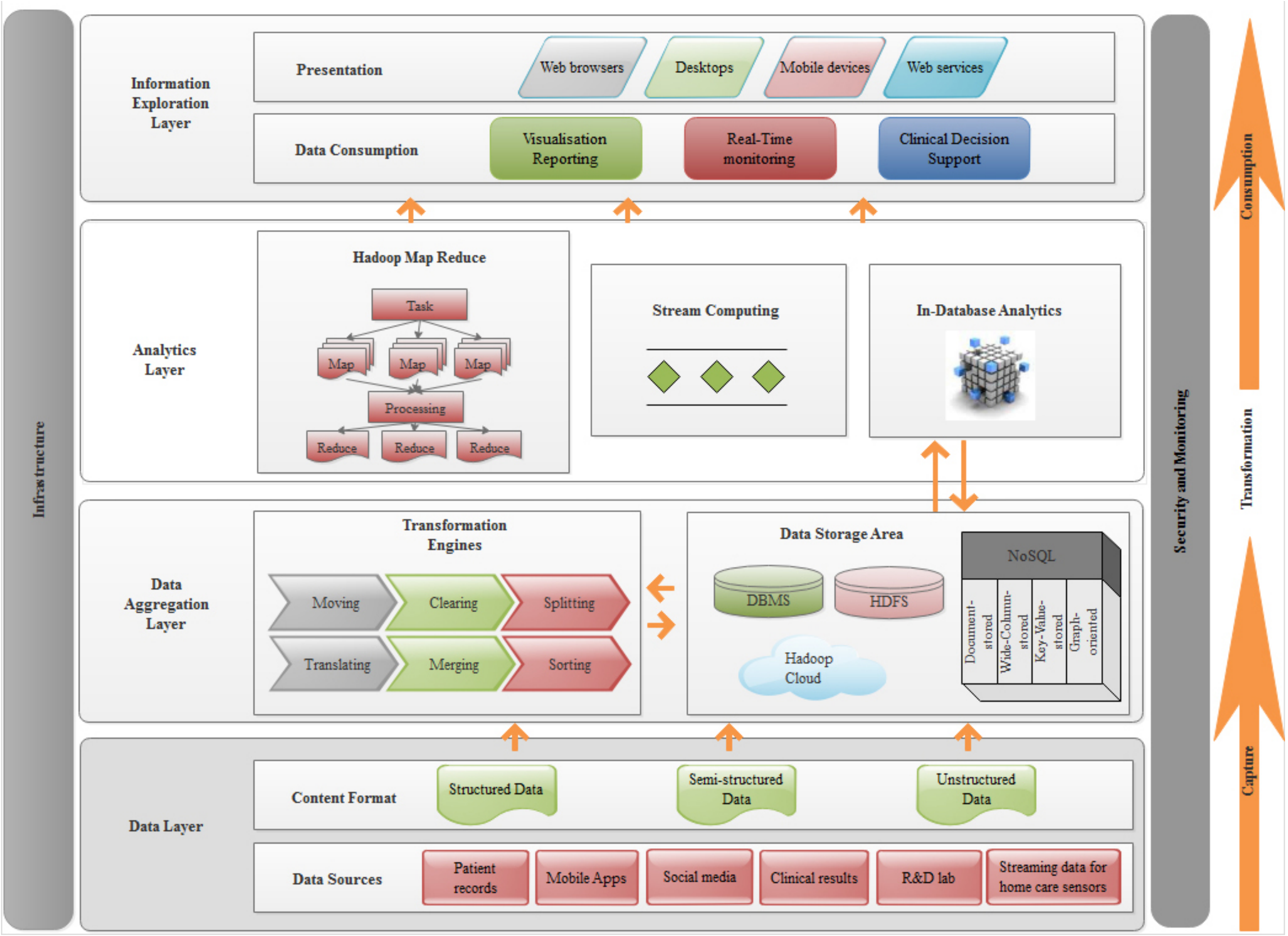


Fig. 2 Big Data architecture for Healthcare (adapted from [30], [31])

database systems based on SQL (e.g. MySQL, Oracle) are no longer appropriate due to their low scalability, extensibility and velocity [32]. They must thus be replaced or used together with distributed, non-relational NoSQL databases which were built for large-scale data storage and parallel processing over numerous servers [13]. Four database models are available [33]: Document-stored (e.g. MongoDB, CauchDB), Wide Column-stored (e.g. HBase, Cassandra, DynamoDB), Key-Value stored (e.g. Redis, Riak) and Graph-oriented (e.g. Neo4J).

3. ***The analytics layer***: processes and analyses the different types of data. The following types of data analysis can be used [30]:
   - *Hadoop Map/Reduce*: can process huge datasets in batch form economically, and analyze both structured and unstructured data in a massively parallel processing environment;
   - *Stream computing*: can perform high performance stream data processing in real time or near real time, and allow for the real- time analysis of data such that events can be identified and responded to immediately. It can, for example, help prevent healthcare fraud by identifying illegal actions leading to it in real time and generating alarms;
   - *In-database analytics*: facilitates processing within the data warehouse by using data mining on an analytic platform. It allows for high-speed parallel processing, but its results are not in real time and thus only a static prediction may be possible. Other advantages are its scalability, security, confidentiality (very relevant in healthcare) and availability of optimization features. In-database analytics is useful for preventive healthcare and pharmaceutical management.
4. ***The information exploration layer*** – can generate different types of outputs [30]:
   - *Visualization reports*: allow for different data visualizations which can improve healthcare daily practice and decision making;
   - *Real-time monitoring of information* (e.g. of physiological readings, alerts, data

navigation, etc.): can be used to assess a patient's health and help avoid exacerbations;

- *Clinical decision support*: the analysis of data about patients and their care can yield new discoveries which support evidence based medicine.

# 5 Applications of Big Data in Healthcare in Romania

The 2014 National Strategy on Digital Agenda for Romania ([34]) stressed on the importance of BDA for helping to re-shape current governmental operations in order to make them more efficient and effective. In what concerns healthcare, BDA was promoted for the better management of funding, the reduction of overpayments, the prevention of fraud and the assurance that benefits are provided to eligible citizens, all of which being great challenges for healthcare organizations. Moreover, the adoption of BDA was proposed as a way to improve the flexibility of healthcare in Romania in the following ways [34]:

- Adjusting the supply for drugs and hospital resources according to insights regarding demand at different times
- Categorizing patients according to insights on their demographic characteristics, and creating personalized services for them
- Enabling the more efficient allocation of patients to healthcare centers, in light of findings on these centers' previous availability.

ICT for Health is currently one of the main Strategic Objectives for which Romania participates to EU funding [35]. The improvement of health and health services is an area where BD projects could bring great benefits for Romania.

Some BD based Romanian healthcare systems are briefly described below [36]:

1. *The SIUI Integrated EHR (Electronic Health Record) System*: an information system owned and managed by the National Health Insurance House (CNAS) which integrates health insurance houses (CJAS) from different counties with healthcare and pharmaceutical providers. It has been in use since 2008;
2. *The SIVMED Integrated EHR System*: a complex solution providing both medical and administrative intra-hospital services to healthcare centers;
3. *The SIPE Electronic Prescription Information System*: an extension of the SIUI system which supports the introduction of electronic prescribing. It has been in use since 2012;
4. *Armonia®sanita*: a hospital information system available on the cloud, supporting numerous hospital services including appointments, hospitalizations and the management of labs, pharmacies and surgery;
5. *Integrated Emergency Service System*: a system which coordinates and improves the efficiency of emergency services at a national level. It helps handle 1,400,000 emergency calls each month.

# 6 Conclusion

BD and BDA can help make healthcare more effective and efficient. They can be used for a range of operations, from disease management and prevention to medical research, and lead to insights which support healthcare providers in making more timely and informed decisions about the population they are managing. Cloud computing offers the storage and processing power necessary for making this possible, and so the combination of BD, BDA and cloud computing is often promoted as a recipe for success.

However, challenges also exist for exploring the potential of BD in healthcare, the most notable being the difficulty of analyzing huge volumes of data to obtain accurate results in a timely fashion, the need for standardization, interoperability, security, privacy, and expertise and funding for developing the BD infrastructure and integrating the already available datasets. New statistical methods, tools, and technology approaches such as cloud computing and security technologies must be explored. Moreover, healthcare organizations should invest in training staff on the use of BD.

In Romania, there is an increased strategic interest in BD and BDA for making healthcare more efficient and flexible and dealing with financial challenges.

*References:*

[1] W. Raghupathi, V. Raghupathi, Big data analytics in healthcare: promise and potential, *Health Information Science and Systems*, 2014 Feb 7, 2:3, pp. 1-10.

[2] B. Feldman, E.M. Martin, T. Skotnes, *Big Data in Healthcare Hype and Hope*, October 2012. Dr. Bonnie 360, 2012. http://www.west-info.eu/files/big-data-inhealthcare

[3] S. LaValle, E. Lesser, R. Shockley, M.S. Hopkins, N. Kruschwitz, Big data, analytics and the path from insights to value, *MIT Sloan Manag Rev*, 2011, Vol. 52, pp.20–32.
[4] I.A.T. Hashem, I. Yaqoob, N.B. Anuar, S. Mokhtar, A. Gani, S.U. Khan, The rise of "big data" on cloud computing: Review and open research issues, *Information Systems*, 2015 Jan 31, Vol. 47, pp. 98-115.
[5] L. Chih-Wei, H. Chih-Ming, C. Chih-Hung, Y. Chao-Tung, An Improvement to Data Service in Cloud Computing with Content Sensitive Transaction Analysis and Adaptation, *Computer Software and Applications Conference Workshops (COMPSACW)*, 2013 IEEE 37th Annual, 2013, pp.463–468.
[6] P. Piletic, *How is cloud and big data changing healthcare in 2016*, 2016, http://bigdata-madesimple.com/how-is-cloud-and-big-data-changing-healthcare-in-2016/.
[7] G. Pres, Big Data Definitions: What's Yours?, http://www.forbes.com/sites/gilpress/2014/09/03/12-big-data-definitions-whats-yours/#1b7816f321a9
[8] Oxford English Dictionary, http://www.oed.com/view/Entry/18833#eid301162178
[9] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, A.H. Byers, *Big data: The next frontier for innovation, competition, and productivity*, McKinsey Global Institute, 2011
[10] V. Mayer-Schönberger, K. Cukier, *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt, 2013
[11] Frost & Sullivan, *Drowning in big data? Reducing information technology complexities and costs for healthcare organizations*, 2012, Retrieved from: https://www.emc.com/collateral/analyst-reports/frost-sullivan-reducing-information-technology-complexities-ar.pdf.
[12] E. Mcnulty, Understanding Big Data: The Seven V's, *Dataconomy*, May 22, 2014, Retrieved from: http://dataconomy.com/seven -vs-big-data/.
[13] A. Alexandru, C.A. Alexandru, D. Coardos, E. Tudora, Big Data – Concepts, Technologies and Applications in Public Sector, *Proceedings of 18th International Conference on Applied Computer Science and Engineering ICACSE 2016*, Prague, Czech Republic, October 6-7, 2016, 7 pages (under publication).
[14] B. Hamilton, Big data is the Future of Healthcare, *Cognizant 20-20 insights*, September 2012, http://www.slideshare.net/cognizant/big-data-is-the-future-of-healthcare.
[15] HIMSS, *Clinical & business intelligence: an analytics executive review*, 2013.
[16] W. Raghupathi, V. Raghupathi, An overview of health analytics. *Journal of Health & Medical Informatics*, 2013 Oct 7.
[17] IBM, The value of analytics in healthcare, IBM Global Business Services, 2012.
[18] HealthCatalyst, Predictive Analytics Solutions, https://www.healthcatalyst.com/predictive-analytics
[19] D. Adamson, *Big Data in Healthcare Made Simple: Where It Stands Today and Where It's Going*, https://www.healthcatalyst.com/big-data-in-healthcare-made-simple
[20] N. Sultan, Cloud Computing for education: A new dawn?, *International Journal of Information Management*, Vol. 30, No. 2, 2010, pp. 101-182.
[21] P. Mell, T. Grance, The NIST definition of cloud computing (draft), NIST Spec. Publ. 800 (2011) 7.
[22] L. Badger, et al., *US Government Cloud Computing Technology Roadmap -Volume II, Useful Information for Cloud Adopters*, NIST Special Publication 500-293, October 2014, http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.500-293.pdf.
[23] CSCC, *Impact of Cloud Computing on Healthcare*, November, 2012, pp. 1-18.
[24] E.C., The Use of Big Data in Public Health Policy and Research, Background information document, Brussels, 29 August 2014.
[25] D. Boyd, K. Crawford, Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon, Information, *Communication & Society*, 2012, Vol. 15, No. 5, pp. 662-679.
[26] Schumpeter, *Little things that mean a lot*. The Economist, 2014.
[27] D.M. Lazer, et al., The parable of Google Flu: Traps in big data analysis, *Science*, 2014, Vol. 343, No. 6176, pp. 1203-1205.
[28] K. Roney, If Interoperability is the Future of Healthcare, What's the Delay?, *Becker's Hospital Review*, 2012.
[29] N. Hashmi, The challenges of implementing big data analytics in healthcare, *TechTarget*, HeathIT, 2013, http://searchhealthit.techtarget.com/tip/The-challenges-of-implementing-big-data-analytics-in-healthcare.
[30] Y. Wang, L. Kung, C. Ting, T.A. Byrd, Beyond a technical perspective: understanding big data capabilities in health care, *System Sciences*

*(HICSS), 2015* 48th Hawaii International Conference on 2015 Jan 5, pp. 3044-3053, IEEE.
[31] A. Alexandru, D. Coardoș, E. Tudora, Using Big Data in the Government Sector, *Proceedings of The 15th International Conference on Informatics in Economy (IE 2016), Education, Research & Business Technologies*
[32] P. Sadalage, NoSQL Databases: An *Overview*, Retrived from: https://www.thoughtworks.com/insights/blog/nosql-databases-overview.
[33] D. Carstoiu, *Distributed database systems*, Conspress Publishing House, 75 pages, 2013.
[34] National Strategy for the Digital Agenda for Romania 2014-2020, *Monitorul Oficial al României*, Partea I, No. 340 bis/19.V.2015, Retrived from: http://www.mcsi.ro/Transparenta-decizionala/Proiecte-2014/Digital-Agenda-Strategy-for-Romania,-8-september-2
[35] EU, Digital Economy and Society Index 2016, Country Profile Romania, https://ec.europa.eu/digital-single-market/en/scoreboard/Romania.
[36] A. Alexandru, D. Coardos, Big Data in Healthcare and Medical Applications in Romania, *2016 IEEE International Conference on Automation, Quality and Testing, Robotics, THETA 20th edition*, 19th -21st May, Cluj-Napoca, Romania, 6 pages, ISBN: 978-606-737-160-4.