

Chapter 2

Big Data in Healthcare

Yiannis Koumpouros

Technological Educational Institute of Athens, Greece

ABSTRACT

The era of open data in healthcare is under way. The progress in technologies along with their adoption by the healthcare providers and the maturity of the citizens has brought the healthcare industry to the tipping point. An unprecedented amount of healthcare data is being generated today. This data comes from researchers, healthcare professionals and organizations, and patients. If we can harness this data, it can help us improve our understanding of disease and pinpoint new and improved therapies more efficiently than ever before. Big Data technologies are coming to market in a rapid way. The challenges, however, are still there due to fragmented systems and databases, semantic differences, legal barriers, and others. The hidden and unexploited knowledge is hindered by these barriers. The big data revolution promises a solution both to this situation, as well as to act as a catalyst to the viability of the healthcare systems. This is supported by the numerous efforts and explored in this chapter.

INTRODUCTION

Healthcare is the industry that first promotes and adopts as well the Information and Communication Technologies (ICTs) innovations (Koumpouros, 2012). The individualities and characteristics of healthcare develop a unique area that requires special care and attention in all aspects. Security, privacy, quality are only some of the concerns raised and studied for many years. Another characteristic of the healthcare domain is the production of big volumes of data and information produced every second. A patient, every time that passes through i.e. a hospital, leaves thousands of “footprints” in his/her visit. These “footprints” are either demographic data, historical data, illness

related information, new or older exams, other relevant information, etc. The patient’s pathway is therefore filled with a significant amount of data that have to be treated accordingly in order to have the desired outcome. On the other hand, the healthcare providers (physicians, hospitals, diagnostic centers, etc.) have to collect and store all this information for each and every patient in a proper manner and be able to reproduce and exploit it for achieving the best possible result. And this is only the one side of the coin. The same provider has also to collect and relate to each patient the appropriate data regarding his/her healthcare consumables, pharmaceutical therapy, other special materials used, etc. Thinking of the number of patients multiplied by the volume of

DOI: 10.4018/978-1-4666-6339-8.ch002

data, times the number of visits we can just imagine the huge volume of information exchanged in the healthcare industry.

Nowadays, living in the knowledge society, data and information are precious. Most of the newly established well-known successful businesses are based on these (i.e. Facebook, Google, etc.). Innovation and new business models arise, constantly utilizing data-driven discovery. One of the main needs is automatic knowledge extraction and exploitation of the produced data and information. This is true in the physical, biological and cyber world. The term Big Data is widely used recently in order to explain most of the above mentioned issues. There is no single formal definition of Big Data. However, most of them seem to coalesce around the “four Vs”:

- **Volume:** The massive volumes of data produced, collected and exchanged.
- **Velocity:** The need for continuous and repeated analysis in real time.
- **Variety:** The many different forms of data that is difficult to integrate.
- **Veracity:** The issues related to uncertainty and trust.

A sea of data has been formed that becomes bigger and bigger every single day. The major problem faced in all industries and by all users is to “fish” (find and collect) the appropriate data out of this deep blue data-sea. A survival scenario is therefore needed to be rescued by this data flood. There is an immense opportunity therefore for forward-thinking leaders and researchers in every industry. To fully leverage the potential that exists within this storm of structured and unstructured data, organizations must quickly evaluate their business, needs and structure, improve and optimize them targeting the right ICTs. Big data are there, but we have to find ways to exploit them and produce the desired valuable knowledge.

Healthcare is one of the leading sectors in the Big Data era. Producing enormous amounts of

sensitive data every minute, healthcare has to find its way in order to utilize the generated valuable information. To this end, several methodologies and technologies have been introduced in order to deal with the Big Data. The need to derive predictive insights from this data for improved business operations and decision support is what drives Big Data analytics. This allows for predictive, contextual, agile, real-time exchange of information across silos. In the healthcare industry (Frost & Sullivan, 2012) introduced Advanced Health Analytics.

Regarding the healthcare’s data we can adopt the aforementioned definition as follows:

- **Volume:** The volume of structured and unstructured health data is enormous, and becomes many times bigger if we could integrate all the data sources from the different healthcare providers.
- **Velocity:** There is an imperative need for real-time analysis of the generated data instead of working on key performance indicators. New algorithms and advanced analytics are therefore needed to conclude invaluable information.
- **Variety:** Today, health data comes from numerous heterogeneous sources and is extremely difficult to aggregate, compare or utilize them accordingly.
- **Veracity:** The unique characteristics of the data produced in the healthcare industry make them really sensitive. Data quality issues are a particular concern in healthcare. The relevance and meaningfulness of the data and their use for decision making are some of the concepts lying under this characteristic. The main point of veracity is whether or not we can trust the collected data as well as its source to proceed further.

The objective of the chapter is to provide a clear view and understanding of the Big Data evolution that emerges nowadays in the healthcare sector. It

22 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/big-data-in-healthcare/116206

Related Content

Workarounds: Risk or Benevolence for Patient Safety?

Nada Nadhrah and Vaughan Michell (2015). *Healthcare Administration: Concepts, Methodologies, Tools, and Applications* (pp. 285-322).

www.irma-international.org/chapter/workarounds/116221

Edge Detection on Light Field Images: Evaluation of Retinal Blood Vessels Detection on a Simulated Light Field Fundus Photography

Yessaadi Sabrina and Laskri Mohamed Tayeb (2019). *Intelligent Systems for Healthcare Management and Delivery* (pp. 174-197).

www.irma-international.org/chapter/edge-detection-on-light-field-images/218119

Leading the Technological Innovation in Healthcare Systems: The Telematic Medicine Approach

Vincenzo Gullà (2015). *Healthcare Administration: Concepts, Methodologies, Tools, and Applications* (pp. 1549-1569).

www.irma-international.org/chapter/leading-the-technological-innovation-in-healthcare-systems/116292

Role of Quantum Computing for Healthcare

Harpreet Kaur Channi and Chiranjeevi Lal Chowdhary (2023). *Handbook of Research on Quality and Competitiveness in the Healthcare Services Sector* (pp. 84-96).

www.irma-international.org/chapter/role-of-quantum-computing-for-healthcare/320843

Medical Tourism for Cosmetic Procedures: Credibility Assessment of Service Providers via Online Health Forums

Makoto Nakayama, Charlie C. Chen and Peter Ratham (2018). *Handbook of Research on Emerging Perspectives on Healthcare Information Systems and Informatics* (pp. 1-17).

www.irma-international.org/chapter/medical-tourism-for-cosmetic-procedures/205141