Schriften zur Gesundheitsökonomie, Vol. 84, Bayreuth: P.C.O.-Verlag.

Big Data: A Panacea to the Health Care System's Challenges?

Valmir Hajdari

In the last two decades, "Big Data," the analytics of enormous amounts of data files, has had an exponentially growing impact on the economy. Software solutions are being heavily researched, although so far, both technological and scientific advances are not developed enough to create programs able to accommodate the necessary sample size and thus, truly reliable information. There are countless possible applications for "Big Data" analyses, especially in industries that heavily rely on statistical data sets, such as the health care sector. Due to the current lack of software solutions capable of coping with the large data sets, today's analysis and usage of "Big Data" is limited. Nonetheless, today's software is still able to sort through unstructured patient data much faster than any manual process could. The results of first programs in the US have shown great potential in solving problems in health care. These improved financial savings, as well as being able to take preventive measures concerning certain diseases, through to the data analysis of a whole society, and even saving lives by accelerating diagnostic procedures. Although there are certain concerns about privacy and security, this paper shows that in the future "Big Data" analysis will tackle numerous heath care related problems and increase the efficiency and efficacy of public health.

Contents

1 Big Data: Could the Use of Big Data help to solve problems of the Health care?

The first use of the term “Big Data” is documented in a 1997 paper by NASA scientists. They described an issue which concerned the problem of the visualization of large data sets (Cox, 1997, p. 235). “Big data” as it is understood today, is defined as large electronic records which are so large and complex that they are almost impossible to analyze with conventional software programs (Kayyali, 2014, p. 2). A report from 2012, which was presented to the US Congress, defines “Big Data” as: "large quantities of high-speed, complex, and variable data that require advanced techniques and technologies to facilitate the capture, storage, distribution, management and analysis of data" (IHTT, 2014). The data management tools developed so far are inadequately suited for managing these records (Frost & Sullivan, 2015). “Big Data” is not only unique because of the volume, but also because of its variety and velocity (Frost & Sullivan, 2015).

Looking at health care, we can see that the health care system historically has collected and stored a very large amount of data (Kayyali, 2014, p. 5). The numerous records and measurements of patient and hospital information can be defined as “Big Data”. It includes data such as medical imaging, clinical decisions, doctor's letters, laboratory tests, prescription and insurance data, as well as other patient records. However, some patient-specific pieces of information, such as social media contributions on social networks, can also be part of Big Data (Bian et al., 2012, p. 26). A study published by McKinsey sees health care as one of the five sectors with the greatest potential for “Big Data” applications (Kayyali, 2014, p. 2). Compared to the past when the data was mostly available in printed form, the trend nowadays is heading further towards the digitization of these existing data sets. Numerous healthcare challenges (demographic changes, co-morbidities, etc.) and the rising cost of healthcare have led many researchers to believe the vast amount of data is the solution to many problems in health care. “Big Data” can be used to support clinical decision making, disease monitoring of a patient, or to improve public health (Fernandes et al., 2012, pp. 38-42).

Previously published reports show the potential uses of health data. According to one report, in 2011 alone, a quantity of 150 exabytes of health-related data were stored in the US. Due to the advancing digitalization and technological approaches, the amount of data is expected to rise to yottabytes and far beyond this in the foreseeable future (IHTT, 2014). For the healthcare industry, the very large amount of data provides numerous opportunities to improve the status quo of health care. An algorithm that would be able to understand all this data and connect them in the right order could potentially lead to lower costs in health care, and therefore improve public health. The use of “Big

Data" in health care thus holds many significant advantages available (Ikanov, 2014, p. 4-7).

This essay will examine if the use of "Big Data" could help to solve problems of the health care system. It provides an overview of data analysis in the health care sector. It is shown what benefits of Big Data could have for the health care. Afterwards, the characteristic properties of "Big Data" are described in more detail, followed by an architectural framework of big data analytics for healthcare and some examples of "Big Data" in current use.

2 Big Data analytics in Healthcare

Data analysis refers to the examination of data sets which is intended to draw a conclusion from the displayed data (Ikanov, 2014, pp. 4-7). Data analytics is used in all business areas. For example, it may be used to help determine the favorite car color of the population for the respective year, or how the new electronic product launch was adopted on the market. In costly emergency care, data analysis can help select the most efficient treatment and preserve resources. Additionally, data analysis can also provide improvements in preventive care, including predictions of population-related diseases and preparations for preventive treatment (Ohlhorst, 2012). "Big Data" can also be used to provide prognoses for specific epidemiological developments or ensure that specialists are already preparing for a case of flu before the outbreak of a flu epidemic. Medical research, especially, can benefit greatly from "Big Data". The ability to collect all data from a particular case and then filter it according to specific results helps to draw the right medical conclusions about a disease (Deross et al., 2011, pp. 52-67).

Due to the new possibilities such as the Big Data, medical research and the rising cost pressure in medical treatment, the health care sector has had to redefine the reimbursement models. The trend of reimbursement models has gone from the fee-for-service variant to diagnosis-related case groups with flat-rate reimbursement rates. In addition, models such as the "pay-for-performance" are gaining more and more interest from the cost carriers. The "pay for performance" model is primarily about achieving a certain treatment outcome and, in the best case, surpassing this result (Burghard, 2014). The better the treatment outcome at the end of a medical intervention, the higher the service provider is compensated for his or her services. However, the additional financial profit should not be the only reason for the implementation of "Big Data." Improvement of treatment quality as well as treatment result are reason enough for the implementation of "Big data" analyses. (Lavelle, 2011, pp. 23-25).

3 The characteristics of Big Data – Volume, Variety, Velocity

"Big Data" is defined by three main features: volume, velocity, variety. Over time, health-related data is still being created and collected continuously, with a vast amount of data coming together. Although this technical progress has some negative effects, the positive developments and improvements far outweigh the costs. Advances in data management, especially in the context of the visual mix and cloud computing, allow for more effective data collection, storage and evaluation (Feldman et al., 2011). The health-related data can now be accumulated in real-time and at a very high rate. This ever-increasing data represent new challenges. With the increased volume and the change in the variety of old, accumulated and new data, new algorithms for retrieving, analyzing and comparing the data are necessary in order to arrive at a result-oriented treatment decision.

Previously, the majority of health data was statistically comparable files in paper form, such as radiographs and medical reports. Nowadays, a large part of the health data consists of real-time data from patients, which could be analyzed in real-time (Feldman et al., 2011). For example, a particular algorithm could use the data of the operating room monitor in real time to suggest a specific treatment recommendation or a particular medication. The analysis of real-time data may in some cases make the difference between life and death (Ikanov, 2014, pp. 4-7). An example of this could be automatic defibrillators which read the analyses the patient's heart rhythm and suggest treatment.

Through the future use of programs that analyze health data in real time, service providers would be able to reduce mortality and morbidity of the patients with proper and quickly-implemented treatments, and possibly even prevent whole disease outbreaks. An example of this real time data analysis in healthcare in current use is the monitoring of newborns via real time streaming in some Intensive Care Units (ICU), allowing the interception of life threatening infections (IHTT, 2014). The possibility to analyze and evaluate all the collected data of an individual patient would greatly improve the health of the individual but also of the society (Feldman et al., 2011).

3.1 Structured, Unstructured and semi-structured Data

Not only the amount, but also the nature of health data has changed over time. Therefore, future algorithms and analysis methods will have to adapt to new job and economic environmental conditions in order to counteract the speed, the volume and the diversity of the data being collected. In fact, the health data are more of a multimedia format and are often unstructured. The high number of structured, unstructured and semi-structured

data provides the health sector with interesting challenges that must be solved in order to ensure a functioning analysis of health data (Feldman et al., 2011).

Structured data is used to describe data that is easy to query, store, retrieve or even analyze. In addition, structured as well as semi-structured data contain electronic recordings of the instruments and other digital examination data. Further, all the data converted from paper form into electronic data are also added to this group of data. Unstructured data is the common form of data that arise during the care of the patient. These include hand-written notes of care, prescriptions, examination images (MRI, CT), etc. Nowadays, there are also many ways to gather data about one's own health (Manyika et al., 2011). There are smart watches and fitness machines which, if desired, that are capable of recording vital functions and many other health data. Also, on social media platforms, many tools are offered to check out one's own constitution, such as track measuring devices. However, the evaluation of this data is far from being as efficient as desired. An actual synthesis of all data to ensure an evidence-based result is not currently possible. Therefore, to make an interpretable result from this data, it requires more efficient programming capable of evaluation of all collected data, as well as the ability to perform an automatic conversion of unstructured as well as structured data (Kesh et al,. 2007, pp. 39-57).

Consequently, it is stated that the positive potential of "Big Data" lies in the merging of all possible data types. The synthesis of data can help individual patients, as well as an entire population. Today, the large amounts of data support scientific research and development of new processes or pharmaceuticals (Manyika et al., 2011). Enhanced data synthesis could assist the development and approval of improved medicines at cheaper prices and in a faster time period. Also, the prescription of the best possible therapeutic option could be made easier and more effective by the use of "Big Data" (Feldman et al., 2011). "Big Data's" potential to improve efficacy and decease cost is very high throughout healthcare and can lead to the resolution of many current challenges. The scarcity of resources and the consequences of the demographic changes could be mitigated by "Big Data".

3.2 The fourth characteristic of Big Data – Veracity

Beside the three stated characteristics of "Big Data," volume, variety and velocity, scholars have introduced a fourth characteristic feature called veracity. It describes the flawlessness and credibility of "Big Data" sets, as well as their analysis and results. Genuine matching data, especially in health care, is of immense importance (Fernandes et al., 2012, pp. 38-42). Genuine means that the data must be correct and true, and that

the records must be complete. Any decision based on false evidence, in the worst case can decide between life and death. In particular, the unstructured data, which usually arise during the care of the patient, must be checked for their veracity, as this data and its analysis is often very difficult and can produce incorrect results. An example of unstructured data causing poor veracity is the handwriting of a doctor, which can be difficult to read and difficult to interpret, resulting in an end point error. This can lead to a false prescription of medicine or worse, even death. (Feldman et al., 2011).

Veracity is a characteristic which not only is required in health care, but also covers all economic sectors, especially on the payers' side. In many treatment cases the reimbursement of a large amount of money is necessary and therefore the correctness of all data is a prerequisite (including the correct patient, hospital, DRG code, money amount, treatments, and prescriptions) (Fernandes et al., 2012, pp. 38-42). In order to achieve accurate yearly results, it is extremely important that all relevant data correspond correctly to the event.

4 Architectural Framework for Big Data analytics

The conceptual framework for a "Big Data" analysis project is very similar to the widely available health information systems or other analysis programs (Fernandes et al., 2012, pp. 38-42). The basic difference lies in the processing execution. Because analysis of "Big Data" is to be done with very large data sets, the processing of this data is divided into several sections. This method is not simple due to the high volume of data. In addition to the large size, the variety of data is also a reason to divide the data (Deross et al., 2011, pp. 52-67). The processing of such large data sets is a recent development. With the processing of "Big Data," various service providers hope to be able to make better informed health-related decisions, and thereby save resources. The algorithms and models for processing "Big Data" are very similar to the existing analysis tools. The user interfaces used to edit "Big Data", on the other hand, differ completely from the traditional data analysis programs. This is, on one hand, due to the volume and the variety of the data, and on the other hand to the diversity of results displayed. The conventional data analysis tools are generally very user-friendly and transparent. Data analysis tools for "Big Data", conversely, are very complex, require a great deal of skill in the evaluation and visualization of the results, and they are also highly program-intensive. The complexity of the programs results from the high complexity of the data sets (Capgemini, 2013, p. 4). The content of "Big Data" in healthcare can come from internal or external sources. The internal sources are the electronic patient records, clinical reports on treatment and diagnosis, and computerized physician order entry, etc. Data from external sources can come from pharmacies, insurances, labs, etc., and are available in different formats from different locations (Lavelle, 2011, pp. 20-32).

In the first step, the data must be pooled before processing. The data is still raw at this time and must be processed or transformed into system-compatible formats in the next step (Kesh et al., 2007, pp. 39-57). The data can also remain in the raw form and can be processed by commands. Another possibility to edit "Big Data" is the so called "data warehouse." With a data warehouse, data from different sources are aggregated and adapted for processing. After extracting, transforming and uploading the data, these are cleaned up and prepared for analysis. Depending on whether the data is structured or unstructured, several data records can also be entered into the data analysis programs for processing (Ikanov, 2014, p. 5).

In the next step, when creating a conceptual framework, further decisions must be made regarding individually relevant functions. It is, for example, necessary to decide which possibilities for editing or inputting information must be available. It must also be decided on possible transmission functions to other carriers or devices. Of course, decisions on the tool selection for processing the data records are still required.

An existing platform for data processing of "Big Data" is offered by the freely accessible Hadoop program. This platform has the possibility to edit very large records by distributing the records to different servers which work parallel to the solution of the problem. Finally, the results of all servers are integrated into one (Deross et al., 2011, pp. 52-67). Hadoop has great potential for the processing of large data sets and thus allows companies more alternatives. However, the operation of Hadoop can be very difficult as there are limited professional staff who are able to use this program. Furthermore, the software is difficult to install, maintain, and configure and qualified employees are rare and very costly. Nonetheless, the program is a first step in the right direction to solving future challenges in health care via "Big Data" analysis.

5 Applications of Big Data in health care

Several "Big Data" applications are already in use today. These programs have mainly been implemented to either achieve advanced cost efficiency or to streamline medical processes and treatments. In one example, IBM reported the usage of a software tool by an unnamed, large healthcare provider that collects and analyses unstructured data sets such as physicians notes and reports. Through this process, the data is available quicker which reduces the time sorting through documents and thus saving cost (IBM, 2014).

A different approach is being taken by the Columbia University Medical Center. Their "Big Data" program aims to analyze data files of patients with brain injuries. Correlations that signify serious complications can be diagnosed up to 48 hours faster than by

a manual sorting processes (IBM, 2014). Furthermore “Big Data” programs can have unforeseen positive effects as well. While creating a “Big Data” set, the US insurance company, Kaiser Permanente, found adverse effects concerning a drug used to treat rheumatic pain which has since been withdrawn from the market. A similar process was used to classify groups of persons with increased risks of developing diabetes. “Big Data” software can therefore not only be used to tackle economic issues, but also to further scientific medical discoveries as well (IHTT, 2014).

5.1 The Data Revolution

Nowadays, a large data revolution is in progress, especially in the health care sector. There are many different reasons for this, but above all, this is happening due to the strong increase in the availability of information. Over the past decades, many pharmaceutical companies have gathered years of research and development data into medical databases. Suppliers and payers have also participated in the digitalization process by introducing an electronic patient records, among other things. Gradually, the US federal government and other public interest representatives have opened their medical information in the form of data for research purposes. This includes data from clinical trials or patient-related information.

In the case of “Big Data,” cost-policy considerations are also the most powerful drivers in the demand for big data solutions. After a steady increase over the past 20 years, US health spending has risen to 16.4 percent of GDP (OECD, 2014, p. 219). This value corresponds to nearly $600 billion more than the benchmark for a country of the size and wealth of the USA. In order to counter the resource shortage, many payers like Medicare and Medicaid have changed their remuneration system from fee-for- service model, which rewards the service provider, to high-volume risk sharing agreements that prioritize results. The changes in remuneration have had an impact on pharmaceutical companies, as their medication needs to demonstrate evidence-based benefits. With this new competitive environment, health care professionals have a strong interest in reducing the cost of health care spending through “Big Data” applications (Kayyali, 2014, p. 3).

5.2 Data protection

The use of Big Data in health care requires special data protection guidelines and laws due to the very sensitive information contained in health care data. The current coalition agreement of the Federal Government in Germany already contains keywords such as "opportunities for digitization," "telemedicine," "data protection," "e-care" and "electronic health card." The introduction of the e-health law by the Ministry of Health has

integrated a roadmap for the introduction of a digital infrastructure with very high security standards (Koalitionsvertrag, 2014, pp. 137-144). Furthermore, for the evaluation of health data, a high level of data protection is required in order to obtain public acceptance for the use of Big Data. For example, the start of the big-data project "care data" of the Health and Social Care Information Center (HSCIC) in Great Britain had to be postponed because the confidentiality of the health data to be evaluated, the cancellation in the event of a contradiction from a patient, and the limitation of the data evaluation could not be guaranteed. In addition, the data could also have been used and abused for police and commercial purposes (Striegler, 2014).

Without a public conversation about Big Data and technical and legal solutions regarding data protection, etc., there will be no acceptance for Big Data in society overall or in health care, specifically. Ensuring a high level of acceptance and ensuring a high level of security when using Big Data are two tasks to fulfill.

5.3 Advantages of Big Data for health care

Due to the digitization and the effective use of large data, health care organizations have been able to gain significant advantages (Burghard, 2014). Among other things, diseases can be recognized and treated in earlier stages than before. In addition, numerous questions which have not been answered so far can be solved by the use of "Big Data." Certain developments or results may be predicted and estimated by the analysis of collected data (e.g., Length of stay) and therefore, be better controlled. Potentially medically-unsuitable interventions also can be canceled in advance due to extended data analysis which can provide information about possible complications (including bleeding, clotting, rejection) or the possible risk of infection of a patient (Burghard, 2014). According to McKinsey's business consultancy, "Big Data" and the use of data analysis could save up to $300 billion a year in US health care. Two-thirds of $300 billion would be generated by an eight percent decrease in the current health issues because of a reduction in health expenditures. In addition, according to McKinsey, $165 billion could be saved in the field of clinical operations, and $108 billion could be saved in in the field of research and development (Manyika et al., 2011). In principle, McKinsey states that the use of "Big Data" can reduce the waste in healthcare and, in addition, it can improve efficient treatment of patients. For example, "Big Data" makes it possible to perform a precise comparison between different treatment methods at a clinical operations level, thus providing a response to more cost-effective treatment methods. Additionally, more effective treatment methods can be converted into treatment guidelines with a better result in the treatment of the patient. The use of "Big Data" has a positive impact on the patient and the care provider.

"Big Data" could also be used in research & development. By using different algorithms and statistical methods, the different clinical studies can be carried out more precisely and more effectively. For example, the choice of the trial design or the determination of the participants in a study opens possibilities for more suitable practice through data analysis, since all relevant details are included. Furthermore, the subsequent analysis of the clinical trial is detailed by an algorithm and thus ensures the unwanted effects of an intervention can be quickly recognized (Manyika et al., 2011).

Important long-term changes can be also generated in public health. Through the use of an algorithm, real-time data can perform the monitoring of certain diseases, thereby allowing for a better understanding of the causes and spread of these diseases. In addition, the results obtained can be used to improve public health. Intervention of a possible influenza wave or epidemic may also be made quicker by the analysis of health data and thus protects larger population groups (Manyika et al., 2011). The use of all accessible data also makes the development of interventions or required vaccines faster. The large amounts of data can also be used to determine the needs of patients. This can be used to ensure that the services offered to the patient are increased in a more efficacious manner. A highly developed public health sector can ultimately lead to improved evidence-based medicine (IBM, 2014). By combining data from different areas (clinical data, operational data and financial data), the most efficient treatment and better care can be implemented. The use of "Big Data" and the corresponding data analysis, therefore, has financial and qualitative improvements in health care for the stakeholders as well as the patients.

6 Conclusion

A usable and effective data processing platform that makes it possible to use "Big Data" must include the necessary tools for data analysis if it is to improve both efficiency and clinical outcomes of the health care system. Important factors in such a platform include the user-friendliness of the program, the ability to manipulate the data, the security of the private data, as well as the scalability of the system (Bollier, 2010). In addition, real-time data analysis allows for a quick reaction and important intervention, which is a critical prerequisite for healthcare-associated analysis programs. Similarly, the gap between data collection and data processing should be closed. In addition, special management questions about ownership and governance of and rules surrounding data must also be taken into account. The fact that health data is often not standardized should also be addressed when creating a big data analysis program (IHTT, 2014).

The analysis of large data records, called "Big Data," has the ability to improve the health economy. "The use of Big Data" has the propensity to change the ways of thinking in the health care system by using new technologies and taking into account large data sets. On the one hand, the new findings may reveal new treatment alternatives, and on the other hand, "Big Data" can reduce the consumption of resources and improve the financial efficiency in health care. Therefore, it is only a question of time as to when "Big Data" will used, via appropriate programs, for the benefit of the health care sector. Furthermore, of course, all the issues mentioned above must be addressed. Privacy, in particular, should be taken into account. The violation of privacy would mean a serious loss of public trust in "Big Data" and therefore in healthcare. This could delay the entire implementation of data processing with "Big Data" or prevent it completely.

In general, the use of "Big Data" can help solve numerous health-related problems. Advantages of such applications range from furthering cost savings, or being able to take population based preventive measures, to even saving lives by accelerating diagnostic procedures. Furthermore, precisely planned treatments can be used to avoid wasting resources and double testing. Predictions of different scenarios, such as influenza epidemics, are also made more possible by analysis of "Big Data." For "Big Data" analysis to be advantageous and not misleading, the data sets must be large and extensive enough. This leads to concerns about both privacy and security.
Nonetheless, a widespread availability of "Big Data" software solutions could tackle these issues while solving problems in all areas of the health care system.

References

Bian, J., Topaloglu, U. and Yu, F. (2012), 'Towards Large-Scale Twitter Mining for Drug-Related Adverse Events', *Proceedings of the 2012 International Workshop on Smart Health and Eellbeing - SHB '12.* Maui, Hawaii, USA, 29.10.2012 - 29.10.2012. New York, USA, ACM Press.

Bollier, D. (2010), *The Promise and Peril of Big Data*, Washington, DC, Aspen Institute.

Burghard, C. (2014), *Big Data and Analytics Key to Accountable Care Success,* IDC Health Insights.

Capgemini (2013), *The Deciding Factor: Big Data & Decision Making,* Capgemini [Online]. Available at: http:// www.capgemini.com/thought-leadership/the-deciding-factor-big-datadecision-making (Accessed May 04, 2017).

Cox, M. (1997), *Application-Controlled Demand Paging for Out-of-Core Visualization,* MRJ/NASA Ames Research Center [Online]. Available at: http://delivery.acm.org/10.1145/270000/267068/p235cox.pdf?ip=144.200.17.42&id=267068&acc=ACTIVE%20SERVICE&key=FC66C24E42F07228.59B42124C4E2603D.4D4702B0C3E38B35.4D4702B0C3E38B35&CFID=942547473&CFTOKEN=69237990&acm=1496150945_7f580282c0c126f21747fae7904392ea (Accessed May 04, 2017).

Deroos D., Deutsch T., Eaton C., Lapis G. and Zikopoulos P. (2011), *Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data,* McGraw-Hill: Aspen Institute.

Feldman B., Martin E. and Skotnes T. (2011), *Big Data in Healthcare Hype and Hope* [Online]. Available at: https://www.west-info.eu/it/il-medico-in-tasca/big-data-in-healthcare/ (Accessed October 04, 2017).

Fernandes, L., O'Connor, M. and Weaver, V. (2012), 'Big Data, Bigger Outcomes', *Journal of AHIMA*, vol. 83, no. 10, pp. 38–42.

Frost & Sullivan (2015), *Drowning in Big Data?: Reducing Information Technology Complexities and Costs for Healthcare Organizations,* Frost & Sullivan [Online]. Available at: https://www.emc.com/collateral/analyst-reports/frost-sullivan-reducing-information-technology-complexities-ar.pdf (Accessed May 07, 2017).

IHTT (2013), *Transforming Health Care through Big Data Strategies for Leveraging Big Data in the Health Care Industry,* IHTT [Online]. Available at: http://c4fd63cb482ce6861463-bc6183f1c18e748a49b87a25911a0555.r93.cf2.rackcdn.com/iHT2_BigData_2013.pdf (Accessed October 04, 2017).

Ikanow (n.d.), *Data Analytics for Healthcare: Creating Understanding from Big Data,* Ikanow [Online]. Available at: http://info.ikanow.com/hs-fs/hub/163225/file-18212299-pdf/docs/data-analytics-for-healthcare.pdf (Accessed October 04, 2017).

International Business Machines Corporation - IBM (2012), *IBM Big Data Platform for Healthcare Solutions Brief,* International Business Machines Corporation - IBM [Online]. Available at: http://public.dhe.ibm.com/common/ssi/ecm/en/ims14398usen/IMS14398USEN.PDF (Accessed May 04, 2017).

International Business Machines Corporation - IBM (2013), *Data Driven Healthcare Organizations Use Big Data Analytics for Big Gains,* International Business Machines Corporation - IBM [Online]. Available at: http://www-03.ibm.com/industries/ca/en/healthcare/documents/Data_driven_healthcare_organizations_use_big_data_analytics_for_big_gains.pdf (Accessed April 29, 2017).

Kayyali, B., Knott, D. and van Kuiken, S. (2014), *The Big Data Revolution in US Healthcare,* McKinsey&Company [Online]. Available at: http://healthcare.mckinsey.com/big-data-revolution-us-healthcare (Accessed May 03, 2017).

Koalitionsvertrag (2014), *Deutschlands Zukunft gestalten, zwischen CDU, CSU und SPD,* Die Bundesregierung [Online]. Available at: https://www.bundesregierung.de/Content/DE/_Anlagen/2013/2013-%2012-17-koalitionsvertrag.pdf?%20%20%20blob=publicationFile (Accessed October 04, 2017).

Lavalle, S., Lesser, E., Shockley, R., Hopkins, M. and Kruschwitz, N. (2011), 'Big Data, Analytics and The Path from Insights to Value', *MIT Sloan Managing Review*, vol. 52.

Manyika, J., Chui, M., Brown, B., Buhin, J., Dobbs, R., Roxburgh, C. and Byers, H. (2011), *Big Data: The Next Frontier for Innovation, Competition, and Productivity,* McKinsey Global Institute.

OECD (2016), 'Gesundheitsausgaben', in OECD (ed.) *Die OECD in Zahlen und Fakten 2015-2016* [Online], pp. 218–219. Available at: http://www.oecdilibrary.org/docserver/download/3015045ec094.pdf?expires=149642 0229&id=id&accname=guest&checksum=35EA292ED211115B59BA28981ABE4725 (Accessed October 04, 2017).

Ohlhorst, F. J. (2012), *Big Data Analytics: Turning Big Data into Big Money*, USA, John Wiley & Sons.

Raghupathi, W. and Kesh, S. (2007), 'Interoperable Electronic Health Records Design: Towards a Service-Oriented Architecture', *e-Service Journal*, vol. 5, no. 3, pp. 39–57.

Striegler A. (2014), 'Britische Hausärzte rebellieren gegen Big Data', *Ärzte Zeitung*, 2014 [Online]. Available at: https://www.aerztezeitung.de/politik_gesellschaft/gesundheitspolitik_international/article/855099/glaeserne-patient-britische-hausaerzte-rebellieren-big-data.html (Accessed October 04, 2017).