Rena T. Gasımova

Institute of Information Technology of ANAS, Baku, Azerbaijan

rena.gasimova@gmail.com

PROSPECTS OF THE APPLICATION OF BIG DATA IN ELECTRONIC GOVERNMENT

The use of Big Data in the era of digital technologies may help to analyze some important data for solution of state and municipal problems and to increase the accuracy of future forecasts. The article explores the characteristics, advantages and opportunities of Big Data for the improvement of e-government services through more efficient use of communication and information technology. At the same time, it provides a number of recommendations on the use of Big Data in e-government.

Keyword: *information society, information technology, Big Data, Big Data analytics, e-government, open data, open government, public administration.*

Introduction

Information, first and foremost, is used to manipulate the society and its governance mechanisms, human consciousness and its development potential. Actual and reliable information, presented as open, plays a leading role in the provision of information interaction between the state administration system and its subjects through the modern information and communication technologies (ICT). The concept of the functionality of the public administration system in the information society has also changed. At present, the vast majority of countries is building an e-government, which is mostly associated with the development of new electronic mechanisms based on IT, which ensures transparency and precision.

Today, information society building in Azerbaijan is one of the priorities of the state policy, and consistent, purposeful measures are being taken in this field. The work carried out in the field of ICT and obtained achievements are of great importance for increasing the authority of Azerbaijan among the other countries. Since 2003, Azerbaijan has been consistently implementing state programs for the development of ICT and information society building.

The e-government includes the delivery of the government services and information to the citizens, entrepreneurs, government agencies and other public organizations through ICT and, in particular, the Internet. Along with the performance of the executive power, e-government also supports the performance of the legislative (e-parliament) and the judiciary (e-court) bodies. E-government open data is very important. Thus, the open source constitutes the basis of the development of modern public institutions, including e-government, thereby creating a framework for the government for citizens (G2C) and the implementation of an open government principle. The Open Data Ideology is supported by a number of major states and international organizations [1-3].

E-government allows users to benefit from the offered services anytime and anywhere without going to the government agencies, which leads to the increased citizen satisfaction and trust in public administration. This also enables the government to properly fulfill their duties and responsibilities before the citizens. Consequently, the prevention of corruption, increasing transparency and decline in costs are observed. Therefore, today, e-government is a priority for all developed countries around the world.

Furthermore, it should be taken into consideration that the idea of *Big Data* has recently become popular and the interest in this area has increased. At present, *Big Data* technologies cover all the spheres dealing with large-scale data-processing: e-government, economy, banking, manufacturing, marketing, telecommunications, healthcare, and web analytics. The use of *Big Data* technologies in the e-government enables the state to deliver services to the citizens in real-time, to develop the relationships between the business and industrial sectors, to manage resources,

to expand and more effectively manage the rights and capabilities of the citizens related to the provision of data access. Given these important issues, it is essential to explore *Big Data* opportunities and its existing problems in e-government. The article explores *Big Data* analysis problems in e-government, basics and capabilities of *Big Data* analytics, and studies the experience of some developed countries in this area.

Problems of Big Data analysis in e-government

The development rate of ICTs offers a completely new and effective government system to handle large volumes of data when using the modern systems and resources. Recently, it has been possible to process large volumes of open data. It is known that today, each government is able to obtain a large amount of data from the Internet. Therefore, the use of *Big Data* becomes attractive for every country. This is due to the need for new administrative reforms in the informatization of the government and local self-governance bodies aimed at building an effective governance system.

In September 2008, a special issue of weekly magazine *Nature* dedicated to the Big Data, which has been published for more than 100 years (since 1869), published the article by Professor Clifford L. Big data: How do your data grow? Clifford L. has first used the term *Big Data*. The article attempts to find the answer to the question "How can technology enabling to work with large data affect the future of science? It also shows the role of data in science, especially in e-science. Over the past years, the term has been used in the business sector and academic environment. This field has gained greater interest since 2011 [5, 6].

At present, numerous conferences, symposiums, workshops, and forums on Big Data problem are held and attended by influential international organizations, scientific institutions and giant companies in the field of ICT. They discuss and explore different aspects of large scale data processing. The future research trends are identified in this area based on a number of open questions. At the same time, indicators that show the decline in the expenditures of the companies spent for the use of cloud technologies are highlighted. It should be noted that they mainly focus on the role of public authorities in providing access to open data and their processing tools [7].

However, *Big Data* problem is still insufficiently studied, i.e., the scientific problems of this field are not fully explained from the scientific point of view. The studies provide a comprehensive overview of *Big Data* concept, its essence and classification on different characteristics, its opportunities, problems, and security issues [8].

Big Data problems are mainly related to the real-time processing, search, classification and analysis of rapidly growing data. There are general characteristics to describe *Big Data* regardless of the fields of application. These characteristics include *Big Data* problems in three basic categories: *volume, velocity* and *variety*. In English-language sources it is called "*3V*". The convergence of these parameters defines *Big Data* and distinguishes it from other data [6, 9].

Recently, the generation of large volumes of information, including their rapid growth, has attracted both academic environment and ICT industry. Collecting, managing, storing hundreds of terabytes and exabytes of Big Data and extracting the useful information out of them with existing methods or tools in is a serious problem. Dealing with structured and unstructured information, deeper data mining, and visualization of the results of the analysis are the key issues of *Big Data* analytics. The increase in the volume of data and the need for their real-time analysis have led to the emergence of the trend, i.e., Big Data analytics [6, 10-12].

The characteristics such as large volumes, speeds and complexity that distinguish this trend from other applications require the appropriate technologies. Therefore, today, the giant in Big Data Analysts offer a wide range of specific software and hardware systems: *SAP HANA*, *Oracle Big Data Appliance*, *Oracle Exadata Database Machine*, *Oracle Exalytics Business Intelligence Machine*, *Teradata Extreme Performance Appliance*, *NetApp E-Series Storage Technology*, *IBM Netezza Data Appliance*, *EMC Greenplum*, and *HP Converged Infrastructure*-based *Vertica*

Analytics Platform. In addition, other companies also have software and hardware tools that effectively process large volumes of data. They include *Cloudera, DataStax, Northscale, Splunk, Palantir, Factual, Kognitio, Datameer, TellApart, Paraccel,* and *Hortonworks* [13-15].

The use of software and hardware platforms in Big Data analytics allows not only identifying the data for the solution of many state and municipal issues, but also increasing the accuracy of predictions of the possible results in the future. USA, Japan, Great Britain, Germany, China and some other countries evaluate *Big Data* as a resource of strategic importance.

Obviously, large volumes, first of all, require large-scale storage and distributed processing. Increased demand for multilevel storage systems and the capabilities of cloud technology have increased the interest in *Big Data* analytics. Cloud technology is one of the most successful approaches to large calculations. It is impossible to work with *Big Data* without cloud computing and storing the data in the clouds [6, 8, 16, 17].

Nevertheless, studies show that most countries are far behind in the *Big Data* sphere. In this regard, many organizations use *Big Data* mostly for marketing purposes and are limited to applying the relevant technologies in business analytics. The advantage of Big Data technology is that it analyzes the development prospects of the social service system for the provision of public and municipal services to the citizens, and eliminates the possible problems in a short period of time. Secondly, automated governance of the public administration based on the processing of large data volumes allows the exclusion of human factor to eliminate existing problems as: corruption, unhandled appeals and illegal decision-making. Thirdly, when applying Big Data technologies, their greater effect is predicted to be observed in all sectors dealing with large volumes of data such as production, healthcare, banking, trade, telecommunications, government, and web analytics. [18]

The complexity of using *Big Data* technologies could be a problem years ago when the velocity of information exchange was low, and the warehouses did not provide the storage of all the available information. These problems have now been eliminated by the technology that enables large-scale data exchange, storage, and analysis.

The studies show that *Big Data* is a more promising aspect of the development of not only commercial structures but also of public administration. In the public administration system, *Big Data* may provide open data analysis for optimal scheduling of city bus routes; ensure services to notify the car owners about the plans of the city administration to repair or street cleaning; collect the information on the condition of electricity lines, roads, and water supply systems [19-21].

Another important duty of the state is to ensure security and public order in the country. Experts believe that this task can be accomplished through the complex information systems like Safe City. Modern cities are supplied with the multifunctional systems - transport, telecommunications, electricity and water supply. Big Data can be used more effectively in urban systems with multi-faceted and sophisticated administration systems. The application of these technologies allows the acquisition and processing of various necessary data volumes, and the fast delivery of the obtained results to the relevant state authorities. For example, the analytical capabilities of the processing of *Big Data* volumes, which come from the video cameras of the Russian "Safe City" information-analytical system, help to disclose up to 70% of the total number of murders in Moscow. In the near future, the confidence of the citizens' in the government and their mutual responsibility will be key resource determining the effectiveness of the public administration. The purposeful and sustained use of Big Data, which accumulates the archives and search engines, trading and analytical platforms, various business and state databases, can create favorable conditions for a high level of development of civil society in the country [18, 19].

Capabilities of Big Data Analytics in E-government

The structure of Big Data analytics maximizes the business opportunities to achieve the organizational goals for Big Data analytical systems and their supporting infrastructure. The tools

ensuring the effective solution of *Big Data* analytics should support availability, scalability and high performance. When selecting the tools for the large scale data analysis, it is important to take into account a number of important aspects and important features:
- supporting different types of data;
- packet or real-time data flow processing;
- eliminating minor delay time;
- providing the cheap maintenance through cloud computing.

At the same time, it should be noted that the structure of Big Data analytics for government should support four key components:
1. resource management;
2. data organization and control;
3. open support for analytics;
4. decision and report on visualization.

The experts specify the following steps to build e-government by examining the capabilities of *Big Data* analysts [22, 23]:
- developing a major national e-government plan;
- developing a strategic plan for the management of e-government services;
- defining the needs of all population through their involvement in this process (this allows the citizens to participate in the development of e-government services);
- using the experience of the countries with the highest indicators on e-government development;
- establishing the educational institutions (where employees are encouraged to participate in the development and management of e-government services);
- developing a mechanism for large-scale data management (roles and responsibilities are arranged for the management and decision-making in regard with the e-government services and *Big Data* analytics);
- developing an analytical large-scale database for building an appropriate infrastructure to support long-term investment in the organization's resource and data management, analysis and discovery, decision support and report visualization, including the development of ICT, e-government, *Big Data* analytics and human capital;
- the need to review business-analytics of large data from the national aspect in creating favorable conditions for *Big Data* analytics;
- ensuring the safety of website users, including data security, and developing secure e-government services.

Opportunities and challenges of *Big Data* analytics in electronic government

At present, *Big Data* technology is adopted to serve to the citizens of all countries, to solve national problems (healthcare, terrorism, job creation, natural disasters) and to prevent the threats. *Big Data* analytics tools can provide a solid foundation for the solution of numerous problems. Macroeconomic problems, employment problems, corruption, security, and many other issues can be solved using the opportunities of *Big Data* Analytics. *Big Data* analytics can aim at making justified decisions, economics planning, and policymaking. For example, the United States, Japan, Great Britain, Germany, and China use *Big Data* analytics in e-government services, management, and decision making. This also allows the countries to increase the accuracy of forecasting.

Big Data is characterized not only by its volume, but also by the variety of data and processing velocity. This data includes the real-time transaction information from the citizens, video-data from security cameras, photographs of space research, and the information of the social media interactions.

The collection of large volumes of data in the state and the development of applications is determined for the solution of complex social problems and stimulation of innovations. Increasing

volume of *Big Data* creates both new opportunities and problems for the government. Experts refer these opportunities to the followings [21, 24]:

- applying the business analytics in online information and services provided by the state through the real-time use of information for e-government;
- monitoring and visualizing the effectiveness of the government activities for a decision -making policy based on the dynamic and public participation;
- generating the necessary information for business reorganization and integration of the state into the electronic environment.

E-government has some institutional and technical problems related to the use of *Big Data*. The relevant governance structures should be established to effectively address the institutional problems. These structures have to address the problems such as the common standard of data providing the data interactions, ensuring confidentiality of the data shared and exchanged by the citizen. One of the key issues here is to provide the organizational support by the administration. The technical problems include the lack of experts, the insufficient software development, the integration of different data sources and formats, including data storage and availability.

Today, according to statistics, almost 2.5 billion bytes of data are generated a day, 90% of which are unstructured. Business Analytics offers new opportunities for revealing the *Big Data* value and supporting decision-making in any organization. In countries with a high population, the data access and collection has not been fully solved by the authorities that use these resources for various management purposes. In some countries there are the agencies that are responsible for data collect and control [25, 26].

It should be noted that the state plays a great role in the development of *Big Data*. Since the state has a large amount of data about both individuals and legal entities. Note that the volume of open information in Europe is huge, which creates opportunities for researchers and entrepreneurs. For example, in some European countries, the government agencies have information about the rented houses. The Internet applications have been developed to show where they are available the vacant houses in real-time mode according to this information. This data also includes the information of the police offices that indicates the level of crime in these or other regions, while the data of other services provide information about the surrounding infrastructure. As a result, a person gets the opportunity to choose a more comfortable option. To address such issues, the government agencies try to apply the *Big Data* and *Open Data* concepts. Therefore, not all queried data can be presented for common use. In this regard, the relevant legislative acts should be adopted to regulate the publication of open data, their verification and monitoring. The impact of open data on the EU economy is estimated at billions of dollars. For example, due to the data transparency in Canada, the speculation has been avoided with the tax savings of 3.2 billion USD [27, 28].

Conclusion

However, the following measures may be deemed necessary for effective use of Big Data technologies in e-government:

- the international experience in the field of *Big Data* shall be studied to attain bigger achievements in the development and digitization of the national information resources in the e-government, and the national compliance taking into account the relevant experience shall be ensured;
- *Big Data* technologies practices of the countries with the highest development index of e-government shall be used;
- experience and knowledge exchange with the developed countries using Big Data technologies shall be provided for further effectiveness and transparency of governance;
- policies and strategies shall be implemented for the application of innovative *Big Data* technologies in e-government, taking into account the local conditions, existing tools and future projections.

References

1. Alguliyev R.M., Yusifov F.F. Some actual scientific-theoretical problems and prospects for the formation of an electronic state // Problems of Information Society, 2014, No2, pp. 3-13.
2. European court of auditors, Special Report No 9, http://www.eca.europa.eu
3. European Court of Human Rights, http://www.echr.coe.int/echr
4. Alhomod S.M. and Shafi M.M. Best Practices in E government: A review of Some Innovative Models Proposed in Different Countries // International Journal of Electrical & Computer Sciences, 2012, Vol. 12, No 01, pp. 1–6.
5. Clifford L. Big data: How do your data grow? // Nature, 2008, vol.455, pp. 28–29.
6. Alguliyev R.M., Hacirahimova M.Sh. Big data phenomenon: problems and opportunities // Problems of Information Technology, 2014, No2, pp. 3–16.
7. Imamverdiyev Y.N. Big perspectives and problems of Big Data technologies // Problems of Information society, 2016, No1, pp. 23-34.
8. Gasimova R.T. Big data analytics: existing approaches, problems and solutions // Problems of Information Technology, 2016, No1, pp. 75–93.
9. Alguliyev R.M., Gasimova R.T., Abbaslı R.N. // The Obstacles in Big Data Process, International Journal of Modern Education and Computer Science (IJMECS), 2017, vol. 9, No. 3, p. 28–35.
10. Price R. Volume, velocity and variety: key challenges for mining large volumes of multimedia information / Proceedings of the 7th Australasian Data Mining Conference (AusDM '08), Australia, 2008, vol.87, pp.17–23.
11. Wu X., Zhu X., Wu G., Ding W. Data Mining with Big Data // Journal IEEE Transactions on Knowledge and Data Engineering, 2014, vol.26, no.1, pp. 97–107.
12. InfoSphere Platform: Big Data Analytics, 2013, http://www-01.ibm.com/software
13. Jacobs A. The pathologies of big data // Communications of the ACM. 2009, vol.52. no.8, pp. 36–44.
14. Babu S., Herodotou H. Massively Parallel Databases and MapReduce Systems // Foundations and Trends in Databases, 2013, vol.5, no.1, pp.1–104.
15. Prajapati V. Big Data Analytics with R and Hadoop, Publisher: Packt Publishing Ltd, 2013, pp. 238.
16. Shang W., Jiang Z.M., Hemmati H., Adams B., Hassan A.E. Patrick Martin. Assisting developers of big data analytics applications when deploying on hadoop clouds / Proceedings of the 2013 International Conference on Software Engineering (ICSE '13), NJ, USA, 2013, pp. 402–411.
17. Assunção M.D., Rodrigo N., Bianchi S., Netto Marco A.S., Buyya R. Big Data computing and clouds // Journal of Parallel and Distributed Computing, 2015, vol.79, pp.3–15.
18. Bulgakova E.V., Bulgakov V.G., Akimov V.S. Use of the Big Data in the PublicAdministration System: Perspective, Opportunities, Perspectives // Journal of Juridical Sciences and Practice: The Bulletin of the Nizhny Novgorod MIA, Russia, 2015, No 3 (31), pp. 10–15.
19. Pavlyutenkova M.Yu. Electronic state of Russia - a virtual reality or a new concept of public administration? // Political management and public policy of the XXI century: the State, society and political elites. - M., RAPS; ROSSPEN, 2008, pp. 365–373.
20. Koh C.E., Prybutok V.R., Zhang X., Measuring e-government readiness, Information & Management, vol.45, No 8, 2008, pp. 540–546.
21. Potnis D. D. Measuring e-Governance as an innovation in the public sector, Government Information Quarterly, vol. 27, No 1, 2010, pp. 41–48.
22. Bertot J.C., Choi H. Big data and e-government: issues, policies, and recommendations / Proceedings of the 14th Annual International Conference on Digital Government Research (dg.o '13), NY, USA, 2013, p. 1–10.

23. Harrison T.M. Using big data for digital government research / Proceedings of the 15th Annual International Conference on Digital Government Research (dg.o '14), 2014, NY, USA, p. 309–310.
24. Chen Y.-C., Hsieh T.-C. Big Data for Digital Government: Opportunities, Challenges, and Strategies // International Journal of Public Administration in the Digital Age (IJPADA) 2014, p. 1–14.
25. What is big data? - Bringing big data to the enterprise, 2013, http://www-01.ibm.com
26. The digital universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East. Study report, IDC, December 2012.
    http://www.emc.com/leadership/digital-universe
27. Roy J. Cloud Computing and Gov 2.0: Traditionalism or Transformation across the Canadian Public Sector? // International Journal of Public Administration in the Digital Age (IJPADA) vol.1, No 1, 2014, p. 74–90.
28. Ojo A., Mergel I., Janssen M. Open data to solve societal issues: workshop / Proceedings of the 16th Annual International Conference on Digital Government Research (dg.o '15), NY, USA, 2015, p. 345–347.