

Some indicators of Big Data

Makrufa Hajirahimova¹, Aybeniz Aliyeva²
^{1,2}(*Institute of Information Technology of ANAS, Baku, Azerbaijan*)

Abstract:- In this article the issue of big data is dedicated, the essence of big data is described which is considered key motivating force of the information society. Big data incorporates along the volume of collected information, as well as storage, computing technologies, combines and services. Big data has become one of the rapidest developing spheres of the information technologies and has great potential and value in many fields. Economic potential of big data has been explained briefly. In order to assessment of big data current situation have been analyzed according to some indicators.

Keywords: -Big Data, Big Data indicators, Big Data potential

I. INTRODUCTION

The end of the twentieth century and the beginning of the third millennium is characterized by the emergence of the information society. In this society information has become an important resource as a particular type of goods from the perspective of interests of commercial, private and state. In the past decade the increase of information and communication facilities and ease of accessibility of information has led the increasing of data stream, the volume of data volume has grown in geometric progression every year. Such "explosion" of information has become the serious problem one of the main challenges of the world of the twenty-first century- named Big data.

The rapid growth of digital information and its value has put new demands to available processing technologies of data and withdrawal of useful information. In response to these challenges exposed to a big information flood and big data projects have been processed Google, Oracle, IBM, Microsoft, Amazon, Facebook organization and so on discovered useful information with the help of different algorithms, technologies. Thus, since 2004 processed large scale data in real-time mode and creation of technological platforms such as MapReduce, Hadoop, NoSQL were organized from more depth analytical tools have risen the value of data a bit [1].

Currently, Big data has become one of the driving force of the development of information technology. Already, in the USA, many other western countries, some developed and developing countries this field began to develop rapidly. In the near it is inevitable future economic and political competition between the countries will be based on using of Big data's potential opportunity. In other words, the investigation and application of large scale data will be necessary in order to increase of the competitiveness ability of any country [2]. That's why, the distribution of economic potential of Big data according to available technologies, fields, countries and measuring of value is an actual issue.

II. BIG DATA ECONOMIC POTENTIAL

Growing of volume, speed, diversity and value has resulted to significant changes in transition direction to the model that managed with data in socio-economic activity paradigm. So, Big Data began to play an important role in the creation of social impact, competitive advantage and innovation field[3]. Therefore, Big Data can be regarded as an underlying asset for all sectors, organizations, countries and regions.

Big Data has mainly great value and potential in 5 fields: health care in the United States; public sector administration in the European Union; retail in the United States; global manufacturing; and global personal location data. In 2010 40% of GDP according to the world has been created by these 5 fields[4-5]. In 2020 the volume of useful data is expected to be 16 zetabytes. In society, its citizens and all business sectors it is reality using of this information for creating of value.

Now, big data has become a large volume invested field all the world. Compared to 2013 in 2014 the number of companies which performed projects relate to big data increased by 125%, the market volume 45%. According to valuation of SNS Research organization only in 2015 \$ 40 billion was invested to Big Data. The growth rate of this investment is expected to be 14% in the next 5 years [6-7].

In 2020 in Europe the GDP of big data and open data is expected to increase 1.9 % (206 billion euros) [4]. It is equal to the overall economic growth in one year of all EU countries. In 2020 is anticipated that Big data will organize the contribute of trade to GDP 23%, manufacturing sector 22%, finance and insurance sector 13%, state administration 12%, , information and communication sector 6%, healthcare and social welfare sector 5% [5].

The expansion of using of Big Data will be positive impact on employment and will lead to the creation of many new jobs. According to source of Gartner in 2015 in this field in the accordance with IT at the international level additional 4.4 million jobs will be created. Near to the 1.9 million jobs will be belong to the USA [8-9]. In 2017 are projected 3.75 million jobs will be created in EU [10].

Big Data is already making a significant contribution to the UK economy. The Centre for Economic and Business Research estimated that by 2017 Big Data could contribute £216 billion and generate 58,000 new jobs in the UK and Ireland [11].

III. BIG DATA MEASUREMENT AND INDICATORS

Initially, big data have occurred as the terms described large-scale data that cannot be processed by conventional methods. In recent years, big data incorporates along the volume of collected information, as well as storage, computing technologies, combines and service.

According to evaluation of Cisco company in 2015 the volume of mobile data traffic was 3.7 exabytes per month. In 2020 it is expected will be 30.6 exabytes [12]. According to research of EMC company in 2020 the volume of data will be increase up to 40 trillion Gbytes, digital data will be 5,200 gigabytes per person in the world. [13].

Although, the new emergency of Big Data it is already stabilized technologies sphere. It has spreaded in many areas of the business and began to play an important role in the development of companies. According to evaluating of IDC company the growth rate of big data is more than 6 times as a whole Big Data market. During the years 2013-2017 the compound annual growth rate (CAGR) of Big Data market is expected 27% and total cost will be \$ 50 billion [14].

Big Data Indicators can be useful in assessing the progress and impact of big data-related growth and performance in the information society (and innovation economy). Generally accepted indicators system for measurement of big data is not still available.

In 2014 initial indicators have been proposed for measurement of big data in Europe market by IDC company. These indicators are as follows [5]:

1. Number of data workers
2. Number of data related companies
3. Revenue of data related companies
4. Data market size
5. Data workers skill gaps
6. Citizen's data

The market value of Big Data is measured in accordance with revenues from the sale of hardware, software and ICT services.

The Innovation Institute at MassTech (Massachusetts Technology Collaborative) has identified eight key indicators that summarize Massachusetts' competitive position in big data and the expansion of the big data ecosystem. These indicators are [15]:

1. Number of big data and data-driven related companies
2. Volume investments in data-driven and big data companies
3. Number of data-driven research centers
4. Number of big data related meetup groups
5. Number of big data related Patents
6. Data science related programs
7. Big Data Projects received federal investment
8. Number of data-related STEM (Science, Technology, Engineering and Mathematics) fields graduate

In scientific sources the number of published documents related to big data can be viewed as an indicator of the big data measurement. Big data research has begun in 2001 with one published document in Scopus [16]. In recent years, published big data documents have increased, in 2014 only in this base the number of published big data documents have reached to 3472. In order to number of these documents the USA, China, Germany, Great Britain and India are the first place [17].

Let us, consider current situation of Big Data market according to following indicators.

3.1. Number of data workers and companies

In order to leverage the potential of Big Data, a key challenge is to ensure the availability of highly and rightly skilled people.

According to investigation result of IDC company in 2014 compared to other countries in USA a number of data workers (10457 people) and a number of data related companies (277821) is quite high. In EU a number of data workers reached 6102 people, a number of data related companies were 243610, in Japan a number of data workers were 3344, a number of data related companies were 95919, in Brasil a number of data

workers were 1031, a number of data related companies were 34840. At the same time the value of data market was €103,935 million in the USA, in EU was €50,454 million pounds, in Brasil was €5,289 million, in Japan was €22,228 million pounds [18].

3.2. Data workers skill gaps

The digital skills gap is one of the most urgent policy challenges facing the developed, and also developing countries. Unlike conventional analytics, mining big data requires an extremely diverse set of skills like data visualization, statistics, machine learning, deep business insights and computer programming. Estimates shows that the US alone faces a shortage of 140,000 to 190,000 people with deep analytical skills, as well as 1.5 million analysts and managers to analyze big data and make decisions based on those findings [11].

The digital skills gap is one of the most critical challenges for UK too. With recent estimates the UK is already losing a potential £2bn per year 28 from unfilled roles requiring digital skills. The techUK survey of tech companies conducted in October 2014 suggested that 93% of tech companies believe the digital skills gap affects their commercial operations and talent acquisition. Tech Partnership have predicted that by 2020 there will be 56,000 Big Data jobs in the UK, a rise from 21,400 in 2013 [19].

3.3. The size of Big Data market according to regions and countries

Big Data is increasingly attracting attention of states which is regarded as a strategic resource in the fields such as science, health, industry, business and etc. Today states have been aimed at allowing the use of large volumes of data to increase the efficiency of decision-making processes, activity of organizations, creation of new services, generation of new ideas and etc [20].

Recently, Big Data was applied in USA companies, now, the other countries of the world interested in this field. In a number of countries are given full support to applying of big data technologies and solving problems in this field at the state level. According to the information of IDC in 2014 according to software, services and facilities of Big Data sphere 45% of the market fell into Europe, Near East, Asia (excluding Japan) and African countries' share.

Also (according to the survey of CIO's) Asia and Pacific Ocean countries have a high rate of development rate in the field of analysis, secure storage and cloud technologies. Only in 2014 in the Asia and Pacific Ocean regions the revenues of big data services amounted 290 billion relate to market [21].

Latin America countries have decided in the 2nd places in order to the volume of investment for the development of Big Data.

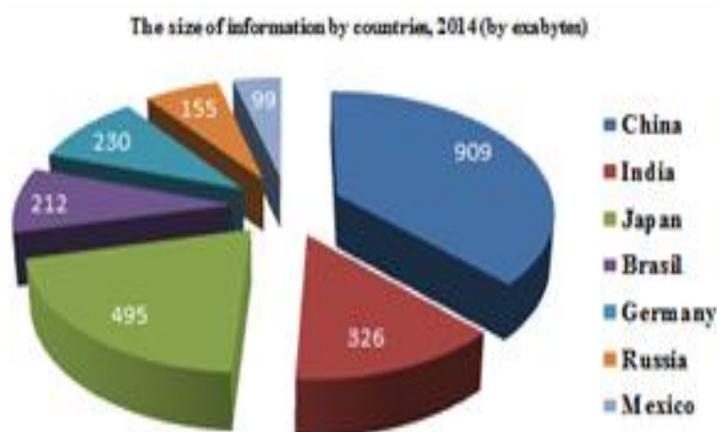
According to evaluation of Cisco company in 2014 the majority of data was accounted in Northern America (1,32zb.) and Asia (1,36 zb.) [21]. Middle East and Africa, Eastern and Central Europe will be the greatest growth rate of big data market (Tabl. 1).

TABLE I. GROWTH RATE OF BIG DATA ACCORDING TO REGIONS

Regions	The volume of Big Data (by zetabytes)		Growth rate (%)
	2014	2020	
Asia	1,36	4,77	24
Northern America	1,32	3,89	20
Western Europe	0,60	1,87	20
Eastern and Central Europe	0,24	1,02	28
Latin America	0,23	0,82	23
Middle East and Africa	0,09	0,70	40

In order to IDC investigation in 2012 the volume of the information was 898 exabytes in USA. In 2020 it is predicted that the volume of information will be 6.6 zetabytes. According to research of EMC company in 2014 the majority of data was accounted China, India, Germany and Japan (Fig. 2) [22]. Digital data was 664,5 gigabytes per person in China, 255,5 gigabytes in India, 3893 gigabytes in Japan, 1045 gigabytes in Brasil, 2839 gigabytes in Germany, 1059 gigabytes in Russia, 789,5 gigabytes in Mexico.

Figure 1. The volume of information by country



According to analysts' prognosis total amount of big data market will be grow result of wide application of big data technologies in developing countries recently. For example, in 2014 the volume of the collected information amounted 40% in developing countries. According to forecast of EMC company in 2017 the structure of Big Data market will be change and in 2020 the share of developing countries will be more than 60%. In China (8060 exabytes), India (2800 exabytes), Brasil (1600 exabytes) and Russia (980 exabytes) the volume of information will be grown rapidly and will account for the majority of data by 2020. [23]. Thus, analysts' forecasts of CIO and EMC company in recent years the developing countries will be active develop market of Big Data technologies.

3.4. Distribution in accordance with technologies

The technologies that used for the collection and processing of Big Data can be divide to 3 places: software (NoSQL, SQL, Infrastructure Software), hardware (compute, storage and networking) and services (professional and cloud services).

The most widely technologies for software contain SQL, NoSQL, MapReduce, Hadoop, SAP HANA.

In order to questionnaires; results of T-Systems 30% of companies that participated in the survey preferred to SAP, HANA, Oracle's "in-memory" platforms, 18% to NoSQL platforms, 15% to Splunk and Dell companies' analytical platforms. 11% of companies preferred to Hadoop/MapReduce products for solving the problems of big data [24].

According to information of IDC the volume of big data world market amounted to 27.3 billion dollars in 2014 and 21.3 billion dollars in 2015. The great part of Big Data revenues were obtained from infrastructure segments (Fig. 2). Their share in total revenue was 51,6%, the share in software infrastructure segments was 23,7%, the share of service segments was 24,7% [9,14].

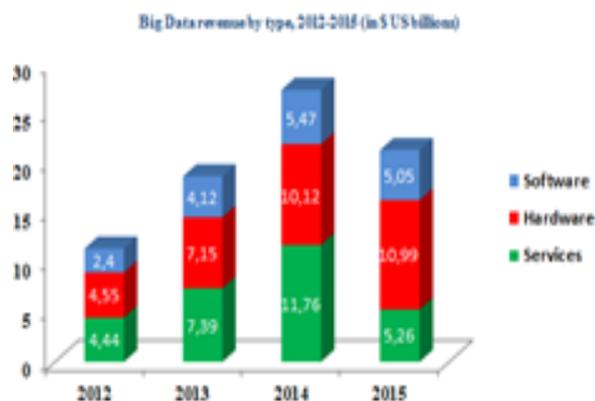


Figure2. Distribution of Big Data according to segments

According to information of IDC company related to application of different fields of the industry, in 2019 the volume of Big Data technologies and services will be increased to \$ 48.6 billion in the world [9,14].

3.5. Distribution according to fields

Approach to Big Data began to rise in order to sources as mobile device, internet, social media, sensors, log file and so on, in most fields (commercial, banking, management, science, healthcare), etc. began to apply widely.

According to estimation of IDC consulting company in 2014 discrete production associate with big data amounted to \$ 2.1 billion, \$ 1.8 billion to bank operations, \$ 1.5 billion to production processes. Fields as securities and investment services (average annual growth rate of 26%), banking (average annual growth rate of 26%) and media (average annual growth rate of 25%) has a higher growth rate [9,14].

According to results of survey of Tech Pro Research company, Big Data has broadly spreaded in the telecommunications sector (58% of requests), IT engineering (surveys, 45%), manufacturing, transportation, financial and state-owned enterprises. In order to the survey healthcare and education are the least spread areas of Big data [25].

According to the forecast of Bain & Company's Insights Analysis company in 2020 the companies of business sector will be the greatest sector, energy sector will be the highest growth rate of Big Data market.

3.6. According to distribution of vendors

Now Big Data has become an integral part of developing of many companies as one of the main directions of IT. According to poll results of Accenture company in 2014 more than 50% of many companies that used Big Data technologies incurred expenditures for Big Data amounted 21% to 30% in the budget [24].

In 2014 Dell company earned \$ 582 million from the sale of big data hardware. At the same time original design manufacturers gained \$5.8 billion from big data hardware [24].

In 2014 according to the services segments revenues of Big Data of IBM company were \$ 624 million [9].

According to the information of Wikibon in 2015 in accordance with the revenues of Big Data services of IBM company, software and hardware segments increased 19% to nearly 2104 million dollars compared to 2014. In SAP company revenues increased 24% to \$ 890 million; in Oracle revenues increased 19% to \$ 745 million. Compare with 2014 in 2015 the revenues of Teradata company rose 50%, the revenues of Splunk company rose 43%, the revenues of Microsoft company rose 37% (Fig. 3) [26].

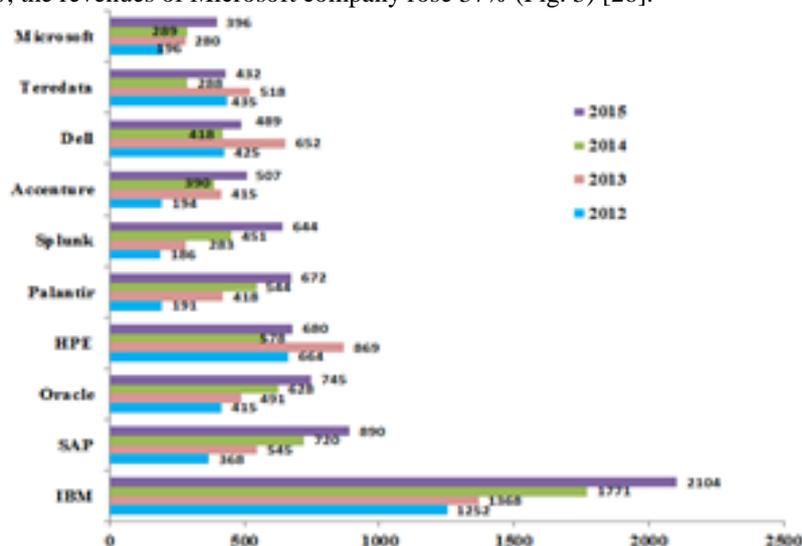


Figure 3. Big data vendors revenue by 2012-2015 (\$US millions)

IV. CONCLUSION

Investigations according to the topic give reason to say that big data has great economic potential. It is no coincidence that big data estimated as a new economic asset at the World Economic Forum in 2012 [3]. There are optimistic forecasts according to growing share in GDP, job creation, increasing investment and so on. Research shows that, indicators system of big data according to the world is not yet completely formed. Therefore, the evaluation of data still remains a problem.

REFERENCES

- [1] R. M. Alguliyev, M. S. Hajirahimova, "Big Data" phenomenon: Challenges and Opportunities, *Problems of Information Technology*, 2, 2014,3-16.
- [2] X. Jina, W. W Benjamin, X, Chenga, Y. Wanga, Significance and Challenges of Big Data Research, *Big Data Research*, 2(2),2015,59-64.

- [3] Big Data, Big Impact: New Possibilities for International Development, 2012, <http://www.weforum.org/reports>
- [4] S. Buchholtz, M. Bukowski, A. Śniegocki, Big and open data in Europe: A growth engine or a missed opportunity? www.microsoft.com/global/eu/Report.pdf
- [5] European Big Data Value Strategic Research & Innovation Agenda - http://www.nessi-europe.eu/Files/SRIA__v099%20v4.pdf
- [6] <https://www.reportbuyer.com/product/2164339/>
- [7] The Big Data Market: 2015 - 2030 - Opportunities, Challenges, Strategies, Industry Verticals and Forecasts, <http://www.researchandmarkets.com/reports/>
- [8] Gartner Says Big Data Creates Big Jobs: 4.4 Million IT Jobs Globally to Support Big Data By 2015, <http://www.gartner.com/newsroom/id/2207915>
- [9] <http://www.statista.com/topics/1464/big-data/>, 2016.
- [10] J. M. Cavanillas, E. Curry, W. Wahlster, The Big Data Value Opportunity, <http://download.springer.com/static/pdf/>
- [11] A.Gupta, Making Big Data Something More than the “Next Big Thing”, The Global Information Technology Report 2014: Rewards and Risks of Big Data, 87-93, http://www3.weforum.org/docs/WEF_GlobalInformationTechnology_Report_2014.pdf
- [12] Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2015–2020 White Paper <http://www.cisco.com/c/en/us/solutions/collateral/service-provider/>
- [13] By 2020, there will be 5,200 GB of data for every person on Earth- <http://www.computerworld.com/article/2493701/data-center/>
- [14] <http://www.idc.com/getdoc.jsp?containerId=prUS40560115>
- [15] Massachusetts Big Data Indicators 2015, <http://massbigdata.org/assets/Uploads/Final-Big-Data-Report-2015.pdf>
- [16] G. Halevi, The Evolution of Big Data as a Research and Scientific Topic Overview of the Literature, *Research Trends, Issue 30*, 2012, <https://www.researchtrends.com/wp-content/>
- [17] R.Aliguliyev, N. Ismayilova, Bibliometric analysis of Big data Research, *Proc. I Republican scientific-practical conference on Big data: opportunities, multidisciplinary problems and perspectives*, Baku, 2016, 58-60.
- [18] F. Lupianez-Villanueva, *Big data in Europe: new environment, new opportunities*, <http://cdn.bdigital.org/PDF/BigDataCongress15/>
- [19] Submission to the inquiry on ‘The Big Data Dilemma’ House of Commons Science and Technology Select Committee, 2015, <https://www.techuk.org/>
- [20] M.S. Hajirahimova, A.S. Aliyeva, Big data strategies of the world countries, *National Supercomputer Forum (NSKF)*, 2015, Russia, Pereslavl.
- [21] Revenue from big data services market in the Asia-Pacific region, <http://www.statista.com/statistics/496266/>
- [22] EMC Digital Universe Study: How Large is the Digital Universe? How Fast is It Growing?, April, 2014, www.infodocket.com/2014/04/16/
- [23] <http://www.emc.com/collateral/analyst-reports/idc-the-digital-universe-in-2020.pdf>
- [24] Analytical overview of the market Big Data, 2015, <https://habrahabr.ru/company/moex/blog/256747/>
- [25] <http://www.techproresearch.com/downloads/the-power-of-big-data-and-iot/#ftag=YHF87e0214>
- [26] http://www.theregister.co.uk/2016/04/08/wikibon_goes_big_on_big_data/.