Journal of Asian Business Strategy

ISSN(e): 2225-4226 ISSN(p): 2309-8295 DOI: 10.18488/journal.1006/2017.7.1/1006.2.23.33 Vol. 7, No. 1, 23-33. © 2017 AESS Publications. All Rights Reserved. URL: www.aessweb.com



BIG DATA MINING AND BUSINESS INTELLIGENCE TRENDS

(Check for updates

Harun Bayer¹ Mustafa Aksogan² Enes Celik³⁺ Adil Kondiloglu⁴ ¹⁴Department of Computer Science, University of İnonu, Turkey. 3Department of Computer Science, University of Kırklareli, Turkey. Email: <u>enescelik@klu.edu.tr</u> *Department of Computer Science, University of Giresun, Turkey.



ABSTRACT

Article History

Received: 02 October 2017 Revised: 16 October 2017 Accepted: 23 October 2017 Published: 8 November 2017

Keywords

Big data Business intelligence Data mining Technology.

The conventional databases are not capable of coping with the high capacity data due to different forms of these data's and fast production speed. In this context, The Big Data structure comes into the scene. The Big Data has been stated as the gold of our age by many authorities. Today, large sizes of data can be analyzed and this led to changes in the lives of people, companies, states, and researchers. The companies develop effective and efficient solutions by analyzing large size of data through big data solutions for their strategic decisions, operational processes, campaign management and marketing techniques. In this research, the introduction has been made to the Big Data architecture, along with daily increasing data mining techniques and methods which will be a solution for accumulating data and current advancements in big data solutions have been addressed. In addition, some well-known companies' tendency to implement business intelligence systems have been examined. The effects of potential threads which are the results of the big data in the business world are analyzed and a couple of suggestions for the future have been presented.

Contribution/ Originality: In this research, the introduction has been made to the Big Data architecture, along with daily increasing data mining techniques and methods which will be a solution for accumulating data and current advancements in big data solutions have been addressed.

1. INTRODUCTION

"In 2020, will exceed the number of stars in the universe in the world of data¹" According to the advancement in technology, the volume of data contained in electronic environment is rapidly

¹ EMC, (2014). <u>http://www.emc.com/leadership/digital-universe/index.htm</u>

growing. Mobile applications, social media tools, sensor operations, electronic systems, online customer relations, cloud storage, web click, mobile traffic, global data returned from performed in the virtual environment in the smart phones, the number of non-structural or structural data is constantly increasing. (Setty and Bakshi, 2013). While the fast data stream in the digital environment enables people to reach information quickly, it becomes hard to differentiate useful information among the giant data that is constantly accumulating.

With the significant increase of the data amount in the world, studies regarding the use of big data have been started. At this point, "big data" term has come into the scene in IT world. In today's world, big data is a big chance. According to the IBM; each day, a data of 2.5 quintillion (2.5x1018) bytes has been created (Grant and Grant, 2017). The 90% of the data of the current day, actually has been created within the last two years. The example sources of data can be stated as receivers for weather information, smart sensors, electronic emails, digital images and videos, instant log records of online shopping websites, cell phone signals etc. These large amounts of data accumulating via numerous sources is always in liquidity. Heading to big data solutions is an inevitable decision for companies in order to analyse this liquid data.

Many companies invest in big data technologies significantly to increase their own value. The companies like Teradata, IBM, Oracle, HP, etc. have been suggesting data-warehouse management to the companies which are in the terabyte level. Today, the data is consolidated in different forms and the structural, non-structural or semi-structural components have been stored as in content form. With big data architecture, companies can analyse a large amount of data and hence, can develop their critical business ideas and manage their strategic planning.

2. BIG DATA

Big Data consists of data which is exponentially increasing in complex and non-structural forms in a never ending speed and these accumulating data has a profile that is far away from resolution option by conventional methods and techniques (Lcia, 2011).

Another explanation for "Big Data" is a structure that analysing the conventional database management with analytic systems and defining the hard or impossible to resolve, rapidly growing, constantly flowing data sets (Partners, 2012). It is hard for the companies and analysts to reach relevant data quickly with this traditional data structures. Recently, in order analyse large amount of data and to provide solutions quickly, "Big Data" is being considered (Devlin, 2013).

2.1. Need of a data warehouse

Today, data is regarded as a treasure which is acquired. It is very important that the data be increased interest in him. Agencies and organizations' need to take advantage of the data in problem solutions has become inevitable. As a result of the increase of interest in the data management, the questions of how to evaluate the data, how it can be functional, and where it will be stored are aroused.

It's very hard to store the data which is flowing through all the time. But if we want to discover useful information from the data and to implement technical analysis, the available data will be needed. Regarding the storing data which is consolidated through many channels, the data warehouses² come into the scene (Nguyen, 2011). Data warehouses are the structures which are accumulated from the many sources that can considered as inputs from decision support systems, along with the recovery of various formats in different giant volumes, they are consisting of relational data resides, capable of making deductions from stored data and replying to the time-oriented, business intelligence applications (Beşikci, 2010).

With the help of data warehouse, converting high-volume, high-speed and various data into a meaningful information and storing them not be possible with conventional methods. With new

technology, storing data and even interpreting the streaming data will be possible (Bahr *et al.*, 2014).

The most important feature in the structure of the data warehouse; for all systems within the company, the only source of information for reports is created. When we export a report through the system previously, we were connecting to multiple different systems and check each system's reporting accuracy. Now thanks to the data warehouse structure, reports can be prepared from the quality and accuracy approved data easily (Inmon, 2013). Corporations need data warehouses to receive accurate decisions from reliable sources. In this way, both the Informatics Department can ensure that clear, accurate and quality information delivered and also the reports exported by different users at different times will not have accuracy problems. Therefore, confidence in the information Department will be increased, analyse and report preparation time will be also decreased.

2.2. With big data and data warehouse

With the increase of products that produces digital data so far, especially in the internet who put what data is not certain and as a noisy data considered as information garbage. Because this noisy, left-over data's storage in databases are hard to be used in analysing and reporting system. But as the technology advances we began to benefit from unstructured waste data as well. Experiencing a production boom digital universe that could be used for decision which are collected from different sources and with data analysis to process all of the architecture led to need of preparation of data stores. To achieve this, the importance of the data warehouse² is highlighted again.

Data warehouse, custom-designed data pool to support the decision-making process. The data in the data warehouse has the following basic properties (Deshmukh and Shelke, 2013).

- Subject-based Data is classified as in logical format based on the main areas of the organization around for example customers, sales or manufactured item.
- The holistic All data that is associated with the subject brought together and be analysed together.
- Time-dependent- For more historical data, it is obtained as detailed form.
- Fixed The data cannot be updated or cannot be changed by users, only can be read.

Rapid developments in information technology, also parallel working processors at the same time, and with the increased use of distributed system in institutions or organizations, the data collection and storage architecture needs led to the Big Data structure; and this stays among the hot topics of current time. In particular, institutions or organizations that wish to increase operational efficiency and to make strategic decisions, they need to be even more involved with the "Big Data" structure.

When making important strategic decision, to enhance operational processes are effectiveness and efficiency, they apply the structure in terms of what we've heard often lately about "Big Data" and "data warehouse" architecture as in Figure 1, with the exact use of the same system, they will be providing a variety of enterprise integration patterns and can put out new ideas (Aydın, 2015).

Depending on the flow of non-structural or structural data, it is seen as possible to increase operational efficiency with the recently developed analysis and data warehouse features integrated into the storage systems. Thanks to the prepared data warehouses, structural or non-structural unlimited data can be used in the analytical process and can create useful query work environments (Göksu and Önder, 2014).

² Data Warehouse concept firstly used by IBM Researchers Barry Devlin and Paul Murhpy (1980) (https://en.wikipedia.org/wiki/Data_warehouse#History)

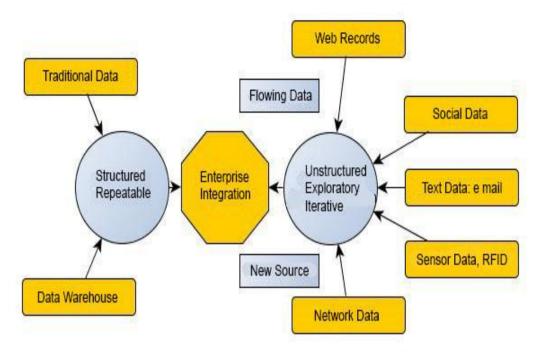


Figure-1. Integrated System on the integration of traditional and new data structures

In the scope of "Big Data", different systems of data storage and analysis are available with distributed architecture. The main systems used; "Hadoop"³, "MapReduce"⁴ and "NoSOL"⁵ systems, within the context of the growing data environment analysis and storage solutions, attract interests of companies that ponders their future. "Hadoop", can perform future-proof highscale data analysis which is constantly growing; the "MapReduce" architecture; is a system briefly analysing by breaking down the data and proving export of requested data. "NoSQL"; manufactures a big data storage solution.

NoSQL, has been developed as an alternative or supportive system to classic relational database management systems that are used for years. NoSQL concept used first in 1998, by Carlo Strozzi in 2009. With the increase of interest in the data, Eric Evans highlighted the NoSQL concept with non-relational, distributed data storage systems in conjunction with the (Corbellini *et al.*, 2017).

Out of a sense of traditional data storage, "NoSQL developed as capable for responding to the needs of the "Big Data" and works with horizontal scaling method (Göksu, 2013). Horizontal scaling means that quick configuration and working with more economical servers with more high-traffic data workload. In this system, unlike the traditional methods, relational tables and columns are not needed. When the system is needed a new space, opening up a new record will be enough. "NoSQL" structure will open a record on behalf of you within the field. We hear a lot the names of Google and Amazon in the world of Informatics and they use "NoSQL" structure and store their giant data for a long time (Lo, 2015).

³ Developed by Hadoop, Apache it is an open source coded software library.

⁽https://hadoop.apache.org)

⁴ MapReduce, Data analysis architecture working on distributed architecture.

⁽www.01.ibm.com/software/data/infosphere/hadoop/mapreduce/)

⁵ NoSQL (2015) SQL and named as more... <u>https://tr.wikipedia.org/wiki/NoSQL (kavram)#Mimarisi</u>

2.3. Data mining

As a result of the increase of large volumes of unstructured data patterns the needs for finding the processing between the data, finding effective results, adjusting according to the needs of fast access, the works regarding this field are increasing. Even small-sized companies can achieve the analysis of the data with traditional methods in their system; thanks to rapidly developing technology, they were forced to use their current data analysis methods and techniques (Han *et al.*, 2011). Presence of the related or unrelated data on the Internet, organizations has to apply Big Data analysis methods and techniques in accordance with new technology. The essence of these techniques actually lies in the concept of "data mining". Data mining is disclosing the hidden and relevant information through the rapidly growing data piles as a result of the set of operations (Roiger and Geatz, 2013).

Data mining, actually works with a similar system that includes the steps of testing in information management, clustering, logical or mathematical queries, summarizing, separation, storage, distribution, access and forwarding. It creates a part of the process of discovering information. Data mining is the step which enables the communication between user and database after the steps of technological evaluation of classification of information obtained, data preprocessing (data extraction, integration, conversion). After this stage, the relationships between the data are reviewed and specific information is obtained.

According to another explanation for data mining states that; "data warehouses will uncover the information previously held undiscovered based on a wide variety of data and according to information discovered, will help company to make the right decision and for their action plan" (Swift, 2001). In this context, based on the big data-driven already, cross-relations between data can be addressed, by disclosing undiscovered patterns and information patterns, it can grant support companies to make strategic and effective decisions.

Data mining make analysis and calculations with the help of several algorithms in the process of knowledge discovery. The most commonly used of these algorithms are as follows:

- 1. Association Rules
- 2. Clustering
- 3. Classification
- 4. Regression

Association rules

It is a data mining rule that is trying to find the relationship of common patterns of the objects within a data set. For example, which products are sold together? Which DNA profiles are sensitive to new drug? with this algorithm questions are proposed.

Clustering

Algorithm which is used for grouping of similar objects in database. For example, determination of different customer profile in markets, while making city planning determination of house prices according to the geographical grouping.

Classification

Data mining algorithm which identifies which class the incoming information belong according to the previously taught, classified database belongs to that assist in example, older women are small, coloured car buys, young men; luxury and a red-coloured car buys ...

Regression

It is the data mining algorithm which determine relationship between more than one variable. As an example can be given as the determination of expenses of potential customers whose income and profession are known.

2.4. Big data mining

As the data-producing digital media trends in an exponential increase, it pushes the electronic measurement system to develop itself. It is common to use data mining in analysing the growing multiple data systems, complex, heterogeneous data as we previously mentioned. But with the growing data, data mining is a particular phase rather than a whole process. In the process of "Big Data " knowledge management, data mining, analysis and other storage technologies can work in an integrated process.

"Big Data" and "data mining" system as in Table1, has some operational and structural differences (Shobana *et al.*, 2015).

(Data Mining)	(Big Data)			
Deals with old big data.	Currently deals with all kinds of information that occurs.			
Data size is small.	Data size is big.			
There are interesting patterns.	Involves the processing of big data sets.			
Data mining does not include Big Data.	In the big data analysis it includes data mining.			

Table-1. The main	differences	between	data	mining	and big	data

Big data mining systems in contrast to the processes of conventional data processing, for data analysis and comparisons many concurrent (parallel) systems are required (Shobana *et al.*, 2015). Therefore, to gain access to at least two types of resource the calculation platform needed: Small-scale data processing tasks would be satisfied with a single hard drive and central processor desktop computer. Actually, a lot of data processing algorithm designed to cope with that kind of tasks. Foe medium-sized data processing tasks, the data is typically larger and probably disorganized. In addition, it cannot be placed into the main memory and does not fit (Wu *et al.*, 2014). In these cases, you can reach solutions possibly by collecting sample and data from different sources, with support for parallel-processor system by using data mining analysis programs.

In Big Data mining, data is well above the capacity of single personal computer. In general, Big data processing task should be performed by data-centric programming language with parallel programming tools. The job of a software components, is to find the most appropriate match in millions of examples in a database and to enable the division of each data task to the smaller tasks which work with one or more calculation node. For example, in Tenesse, located in Oak Ridge National Laboratory the world's most advanced super computer named as Titan has 16-core CPU and have 18.688 node (knuckles) (Wu *et al.*, 2014).

Thanks to the technology developed, the earlier mentioned "MapReduce" architecture that is running simultaneously can work with data mining algorithms and perform various analyses together. In the last decade, some clustering algorithms have become more demanded. But it's hard to do clustering with Big Data. in the architecture of "MapReduce", data sets contained in the nodes can be read quickly. If the number of nodes in a cluster is high, process is that much faster. The operations that will be performed with big data, process can be deployed to multiple processor running in parallel.

In summary, it is hard to process a very large amount of the data with huge sizes by high-equipped servers and it economically costs a lot. In a constantly evolving industry, "MapReduce" architecture which makes analysis simultaneously, quickly and easily, began to be used often (Nandakumar and Yambem, 2014). "MapReduce" architecture do the data break down first, then integrates in a meaningful way. In this way, analyse can be completed quickly.

"Hadoop" architecture is an another architecture used in the process of analysing Big Data. "Hadoop" grouped distribution of large volumes of data between computers. "Hadoop was formed from the clustering structure running under, "Apache⁶" platform which is an open source architecture (Campbell, 2015). Originally designed to search web-related data "Hadoop", as of today, you can make all kinds of establishment and industry analysis and analytical decision for support. For example, it can answer to organization's question of what should be their product investments. Besides, we'll call a person in application through the social media environment. We're looking for names and last names, in a few seconds, thousands of people turning out is a "Hadoop" example of the architecture.

This type of "Big Data" systems that bring together both hardware and software components is unlikely to be supplied without the support of the main industry stakeholders. In fact, for decades the companies work with operational data stored in databases about decisions. With "Big Data" architecture, that they can go further than their database and provide options to go along with less structured data. Such as weblog, social media, e-mail, sensors and photos etc. to obtain useful information. IBM, Oracle, Teradata and the other leading technology companies, adjust all their products to obtain and edit them from many sources by their customers.

3. BUSINESS INTELLIGENCE ORIENTATIONS OF THE BIG DATA

Thanks to the rapidly growing pile of data that companies are looking for meaningful information in these data and with "big data" structure along with the "data warehouse" and "data mining"; which means turning them into the fluid, non-structural or structural information, by storing, you can obtain useful and important information. Now many companies are aware of potential value of big data. Thanks to the "Big Data" companies can respond to the real needs of customers. According to "Gartner"7 as of 2015, the companies' investment quests, if the "Big Data" used in accordance with data warehouse system, the competition in terms of performance and financials, 20% advantage compared to its competitors in the industry can be opened (Microsoft, 2015). In a competitive environment, to survive, to make a difference, to increase the internal dynamics, to develop end user-oriented campaign, to increase the market share of the company, such innovative information management structures are a process that is inevitable for companies. Microsoft has started to develop solutions based on Hadoop architecture on "Big Data" architecture that will support new generation "cloud" or large size data scenarios in order to assist companies' critical decisions which use Big Data architecture on structural and non-structural data. Microsoft, for the "Big Data" solutions, while using parallel data warehouse architecture solutions in both structured and semi-structured data also with "Hadoop" data by setting up an infrastructure, on the same platform, the data is recycled quickly which is extremely important for the company or organizations using these solutions. International Data Corporation (IDC) report says in the world's most admired companies ranking list, the second highest of the EMC company is in a leading position for ten years in the software market of data storage infrastructure software.

⁶ Apache is an open source coded and free web browser program (Wikipedia)

⁷ Gartner Group, Company of research (American) .(<u>https://en.wikipedia.org/wiki/Gartner</u>)

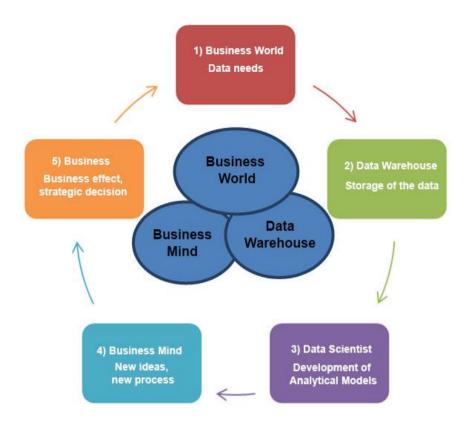


Figure-2. EMC company big data life cycle

EMC companies that provide solutions regarding the data storage process management has a lifecycle with "Big Data" which is presented in Figure 2. In the process developed, the business world, data warehouses and business intelligence processes have been assessed in relation to each other. Life cycle process has a structure that ensures that the data is in the process of continuous innovation. In the scope of the data-related life cycle of the business world, as a top priority of every company, the useful data will be collected according to their own. Collected data are systematically stored in data warehouses. The next stage, the data scientist develops analytical, descriptive and estimation models upon the data which passed through pre-treatment process by comparing the results obtained from different data sets. At this stage, with the help of machine learning algorithms and data mining algorithms, new patterns, templates are discovered and some repeated predictions are tried to be forecasted by using useful prediction models. At the end of this process business intelligence process will be directed (Seker, 2015).

In the process of Business intelligence undiscovered and hardly predictable patterns can be explored through appropriate modelling and critical decisions. Modelling results will be evaluated by the data scientist. According to results discovered, companies can develop new ideas, effective decisions in the business world and can identify strategic targets. For example, which area should the company invest in is a vital information for an investor. Besides, the best-selling and most visited products in the internet is a knowledge which worth golds for a marketer.

In "Big data" solutions, self-renewing "Oracle" integrated a way that could correspond to the needs of companies under one roof by using "big data" for a relational database, data mining structure, the "SQL" cost and performance, "NoSQL" "Hadoop" architecture. Thanks to this integrated system, "Oracle" minimizes the data transactions, guarantees high performance, and also provides data confidentiality. "Oracle" provide "Big Data" solutions for institutions or organizations in the market. For example; estimating the age of customers, prediction of behaviours by using "data mining" along with "Hadoop", "NoSQL" in a way that is integrated with the architecture.

3.1. The danger of big data wave with the business world

With a lot of convenience that comes with virtual world, institutions or organizations in the digital environment applied a business strategy which is "Big Data" oriented. According to the Xerox's research which is an international technology company from U.S., 56% of the companies ' decision making, increase of work efficiency and goal setting processes benefit from Big Data architecture. But 55% of companies overlooks suspiciously to the "Big Data" structure. According to the results, personal data security concerns can affect the development of the "Big Data" (Yaylacı, 2015).

Recently, the European Union Commission data about the safety shows that; studies should be performed for companies' need for the protection of that data in data monitoring, a particular way should be determined according to data map and current flows should be determined. An Antivirus and computer security software Producer "McAfee" performed a security research which states that; the firms only able to detect 35% of vulnerabilities. According to the results of the report, the companies seem to have inadequate security measures. Some companies failed to resolve Big Data. Due to the lack of management "Big Data" processes, they experience trouble in detecting their security flaws as well (Arslan, 2013).

Storage of big amount of data brings main issues such as; loss of encryption, data companies' documentation, individual's privacy, open source software inherits a number of security issues. If we state more specifically; demographic characteristics of individuals, social security numbers, bank account information, passport and driver's license numbers, telephone numbers, fingerprints, job information, employment and financial data, social media accounts are stored in the digital environment and this will create data spies who wish to benefit from these data. The determination of the attacks, prediction of similar patterns; including removing the pattern can be identified. Even exponential encryption techniques and real-time use of "Big Data" integrated management systems used as architectural solution, their success will be still depending on the progress of the technologies in the future.

4. CONCLUSION

Today, the data which companies have turned them into a gold mine, so to speak, and they are like a treasure under a big rock which is waiting to be discovered. In order to reach this valuable treasure, like a master miner, one should make a deep research, dig the right places, make the right analyses and should provide right methods and techniques which will make this valuable treasure unique. Business intelligence solutions with the Big Data are currently in high demand thanks to the companies that aim at coping with these large data stream, in order to survive in this competitive market environment in future. Companies will gain strength in this competitive environment after the correct analyses of social media communications, income statements, the relation between sales and advertisement budgets, office documents, stock market data, investment rates, and even the liquid instant data which is streaming in websites. The data streaming from numerous sources will worth golds if it is processed. However, at the same time, it may bring big problems and security flaws alongside. In this regard, it is understood that companies should make significant investments in Big Data solutions.

Funding: This study received no specific financial support.

Competing Interests: The authors declare that they have no competing interests.

Contributors/Acknowledgement: All authors contributed equally to the conception and design of the study.

REFERENCES

- Arslan, K. (2013). Companies fail at large data security. https://www.technopat.net/2013/06/19/sirketler-big-data-guvenliginde-basarisiz/.
- Aydın, L. (2015). MS parallel data ware. <u>http://www.innova.com.tr/sql-server-data-warehouse.asp</u>.
- Bahr, M., Aydogan, B., Aydin, M., Khodabakhsh, A., An, I., & Ercan, A. O. (2014). Real-time data reconciliation solutions for big data problems observed in oil refineries. In Signal Processing and Communications Applications Conference (SIU), 2014 22nd pp. 1612-1615.
- Beşikci, A. (2010). A decision support system based on data warehouse for a University quality system. Project Fair. Anadolu University Campbell D. Hadoop'un kısa geçmişi.(https://azure.microsoft.com/tr-tr/solutions/hadoop/.
- Campbell, D. (2015). Hadoop'un kısa geçmişi. (<u>https://azure.microsoft.com/tr-tr/solutions/hadoop/</u>.
- Corbellini, A., Mateos, C., Zunino, A., Godoy, D., & Schiaffino, S. (2017). Persisting big-data: The NoSQL landscape. *Information Systems*, 63, 1-23.
- Deshmukh, C., & Shelke, P. (2010). OLAP, OLTP for Decision-making process using data mining and Warehousing. *International Journal of Interactive Communication Systems and Technologies*, (ISSN: 0975-9646).
- Devlin, B. (2013). Big analytics rather than big data. <u>http://www.b-eye-network.com/blogs/devlin/archives/2013/02/</u>.
- EMC (2014). The digital universe of Opportunities. <u>http://www.emc.com/collateral/analyst-reports/idc-digital-universe-2014.pdf</u>.
- Göksu, C., & Önder, A. (2014). Discover your data warehouse competencies with big data. <u>ftp://public.dhe.ibm.com/software/pdf/tr/2_Veri_Ambari_Yetkinliklerinizi_Buyuk_Veri_il</u> <u>e_Genifletin -_Cuneyt_Goksu_Ayhan_Onder.pdf</u>, IBM Big Data Forum. Accessed:14.08.2015
- Göksu, C. (2013). What is this NoSQL. http://datawarehouse.gen.tr/nedir-bu-nosql/.
- Grant, D., & Grant, C. (2017). Missing out: Does masters students' preference for surveys produce sub-optimal research outcomes? *TIRI Conference*. p-54.
- Han, J., Kamber, M., & Pei, J. (2011). Data mining: concepts and techniques. Elsevier.
- Inmon, B. (2013). Big data implementation vs. data warehousing, <u>http://www.b-eye-network.com/view/17017.</u>
- Lcia, (2011). Big data: big opportunities to create business value. <u>http://www.emc.com/microsites/cio/articles/big-data-big-opportunities/LCIA-BigData-Opportunities-Value.pdf</u>.
- Lo, F. (2015). What is Hadoop? What is MapReduce? What is NoSQL?<u>https://datajobs.com/what-is-hadoop-and-nosql.</u>

(2015).

BigData.<u>http://www.microsoft.com/turkiye/kurumsal/datacenter/features.aspx?BigData</u>.

- Nandakumar, A. N., & Yambem, N. (2014). A survey on data mining algorithms on apache hadoop platform. International Journal of Emerging Technology and Advanced Engineering, 4(1), 563-566.
- Nguyen, P. V. (2011). Using data warehouse to support building strategy or forecast business tend. arXiv preprint arXiv:1205.0724.
- NoSQL (2015). Wikipedia, NoSQL Concept. Accessed: 18.10.2015 https://tr.wikipedia.org/wiki/NoSQL (kavram)#Mimarisi.
- Roiger, R. J., & Geatz, M. W. (2013). *Data mining a tutorial-based primer*. Addison Wesley, United State of America, 2003.
- Partners, N. (2012). Big data executive survey: creating a big data environment to accelerate business value. <u>http://newvantage.com/wp-content/uploads/2012/12/NVP-Big-Data-Survey-Accelerate-Business-Value.pdf.</u>
- Setty, K., & Bakhshi, R. (2013). What is big data and what does it have to do with it audit? *ISACA Journal*, 3, 23-25.

Microsoft,

- Shobana, V., Maheswari. S., & Savithri M. (2015). Study on big data with data mining. International Journal of Advanced Research in Computer and Communication Engineering, 4(4), 381-383.
- Swift, R. S. (2001). Accelerating customer relationships, prentice hall PTR a pearson education company. Upper Saddle River, p.71.
- Şeker, S. (2015). The concept of big data and big data life cycles. <u>https://www.youtube.com/watch?v=gi-rM0yRoXQ</u>.
- Yaylacı, S. (2015). Big data is important but it arouses suspicion. (http://www.btnet.com.tr/arastirma/big-data-onemli-ama-kusku-uyandiriyor/1/18647.
- Wu, X., Zhu, X., Wu, G. Q., & Ding, W. (2014). Data mining with big data. Knowledge and data engineering. *IEEE Transactions on*, 26(1), 97-107.

Views and opinions expressed in this article are the views and opinions of the author(s), Journal of Asian Business Strategy shall not be responsible or answerable for any loss, damage or liability etc. caused in relation to/arising out of the use of the content.