International Journal of Advanced Culture Technology Vol.7 No.4 255-259 (2019)
DOI 10.17703/IJACT.2019.7.4.255

IJACT 19-12-30

Utilization and Analysis of Big-data

Soowook Lee[1], and Manyong Han[2]

[1] Associate Professor, Glocal Education Center, Kwangwoon University, 20 Kwangwoon–gil, Nowon-gu, Seoul, 01897, Korea
[2] Professor, Tax & Accounting, Seoil University, 28, Yongmasan-ro 90-gil, Jungnang-gu, Seoul, 02192, Korea
wook@kw.ac.kr[1], myhan@seoil.ac.kr[2]

Abstract

This study reviews the analysis and characteristics of databases from big data and then establishes representational strategy. Thus, analysis has continued for a long time in the quantity and quality of data, and there are changes in the location of data in the social sciences, past trends and the emergence of big data. The introduction of big data is presented as a prototype of new social science and is a useful practical example that empirically shows the need, basis, and direction of analysis through trend prediction services. Big data provides a future perspective as an important foundation for social change within the framework of basic social sciences.

Keywords: *Big Data, Social Science Methodology, Analysis Strategy, Correlation.*

1. INTRODUCTION

One of the most notable concepts in the field of telecommunications is big data. In the past, people affected by analyzing and accessing data more efficiently. Big data means not only a diverse and vast amount of data, but also all relevant processes for analytical technology and valuable use, and is a key element of the Fourth Industrial Revolution, which determines national competitiveness. Big data provides a single version of the information and makes smarter decisions about business processes. However, even though most of the information is unstructured data, most master data management (MDM) solutions only provide structured data. This study will find out the necessity of knowing the concept of big data that may be unfamiliar to us and the knowledge about big data to modern people by studying the field where big data is used in modern society.

It is known that big data, essential for the modern IT society and it can be applied to many areas such as management, finance, economy, data mining and the Internet of Things. Therefore, the study of big data content would help to cope with future society, and it is the process of establishing government policy that can use big data most efficiently.

2. BIG DATA CONCEPTS AND CLASSIFICATIONS

Big data is a technology that makes it easy to share information through portable information and communication devices such as smartphones that establishes information and communication market through

Manuscript received: October 02, 2019 / revised: October 13, 2019 / Accepted: November 22, 2019
Corresponding Author: wook@kw.ac.kr
Tel:+82-2-940-5649, Fax: +82-2-916-4751
Author's affiliation
Associate Professor , Kwangwoon University

the Internet. Traditional big data simply means large amounts of data. Recently, the concept of big data refers to data sets that are difficult to collect, store, search, analyze, and visualize because they are much larger than existing data.

Table 1. Four elements of big data

Division	Main Content
Volume increase	Technological advancement, IT routineization, digital information volume increase Zettabyte entry
Variety increase	Text, Multimedia, Unstructured Data Types
Complexity increase	Unstructured data, Storage differences, Redundancy, Expansion of data types, Data management and processing
Velocity increase	Increased information, Streaming information, Real time increase

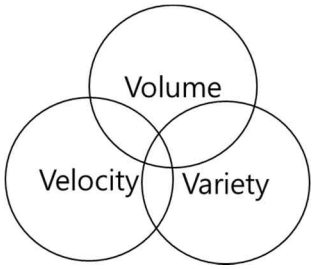

Figure 1. Three major elements of big data

2.1 Big Data Usage Status of Science and Technology Policy

Korea has carried out specific activities for big data policy for the fourth industrial revolution. First, in 2011, the National Informatization Strategy Committee and others announced the implementation of big data to build the smart government. Through this, we established detailed plans for establishing a government data connection and analysis system, improving the legal system, training analytical personnel, and retraining. A detailed plan of data connection and analysis system, improving the legal system, training, etc. will be established. Since 2012, the Korea Communications Commission's 'Big Data Service Activation Plan' and the Big Data Master Plan for the Realization of Smart Countries such as the Ministry of Public Administration and Security have been announced one after another. In addition, as a priority, the government is implementing policies to solve social problems, such as predicting the location and time of crime, early detection of natural disasters, and policies to reduce the incidence of involvement in traffic accidents. It is expected to be more effective in connection with Korea's advanced IT infrastructure.

According to the broadband portal announced by the Economic Cooperation Agency on January 26, 2014, Korea has a subscription rate of 102.12 per 100 people over the Internet. In addition, according to an online statistical survey by Mashable (Mashable, the survey statistics portal site of Mashable) published on IT specialized media, the significant smartphone penetration rate in big data technology development was 73.0%, much higher than the United States (56.4%) and the United Kingdom (62.2%). If Korea's strong IT infrastructure and social problem-solving big data policy are linked, it can provide significant services to solve social problems while providing optimized services to individuals.

2.2 Late Night Buses in Seoul, Korea

The first example of improving public services using big data is the establishment of a late-night bus route in Seoul.

The late-night bus, which operates in areas with a high floating population at night, is a policy that

effectively solves the problems of citizens who have difficult to return home. The process of determining and implementing paths was driven by big data analytics. To collect relevant information, we analyzed about 3 billion data of KT mobile phone data used by citizens from midnight to 5 am in March 2013. The collected information was analyzed by the city's late-night bus route support system. The total population of Seoul is divided by a radius of 1.250 to indicate flow and traffic needs. The re-analysis process yielded optimal routes and dispatch intervals, such as calculating the weight of the population floating near the route.

3. BIG DATA USAGE CASE OF SCIENCE AND TECHNOLOGY POLICY

Many developed countries are now using big data in their policies. In March 2012, the U.S. announced its Big Data Initiative to develop big data technologies, spending more than $ 200 million around the Office of Science and Technology Policy (OSTP). The plan included the development of big data technologies, innovation in education and learning, professional development and program expansion. In addition, the vision of the Big Data Senior Steering Group under the President has continued.

In response to this trend, the US Bus Information System (BIS) established the Data Strategy Committee in March of the same year with the motto of creating value through public information sharing and utilization. In June, we released the Open Data Strategy to help departments identify and leverage access to big data and personal data.

Singapore, led by the Economic Development Bureau, has established a data analytics lab to strengthen government and corporate competitiveness, while the National Security Coordination Department has established a Risk Assessment and Horizon Scanning (RAHS) system that assesses threats based on data. In January 2012, we set up a risk assessment and Horizon Scanning (RAH) to operate it. This allows Singapore to run coastal safety and bird flu simulations and improve national crisis response.

The Japanese government will use the active Japan ICT strategy to fully utilize Japan's growth-based information and communication technology to create new markets and new industries and to create a society where anyone can actively participate in a safe environment regardless of generation or region. In order to realize “ICT Smart Town” by utilizing information communication technology such as sensors, wireless technology, cloud, etc., we carry out various demonstration projects, establish new models, and attempt to spread them.

4. PROBLEMS IN USING BIG DATA IN SCIENCE AND TECHNOLOGY POLICY

4.1 Infringement of privacy and privacy rights

Concerns about privacy and infringement of privacy rights are the most important risks already known in relation to the use of big data, and the main principle of privacy law to protect these rights is the risk of conflict with big data. Privacy and privacy rights violations are very important in that the rights and rights of these rights, as well as constitutional fundamental rights, have direct and indirect relevance to other side effects of the use of big data.

Violation of these rights in accordance with the principle of minimum collection, principles other than the purpose, principles of data quality assurance and transparency, which are defined as basic principles for the protection of personal and personal information, creates tensions that cannot be shared with big data. It's also a matter of urgent review, which can be said to be highly feasible..

4.2 Atrophy effect

If a company collects and analyzes the data that individuals generate in their daily and offline routines, and provides customized products or services, these actions immediately lead to privacy issues and unfair discrimination. However, if such a situation persists for a long time, buying insurance based on a variety of unknown information, depending on the state of health can be denied. The imminent consequences of imagination or realization can have a negative impact on the daily life and social participation of current citizens.

We search and provide other knowledge online because of the concern that tracking and visiting the Internet, networks, and platforms can affect a variety of decisions that are important to an individual's life, such as finding a job in the future. Reluctant to express various views. Collecting, analyzing, and tracking information about people and activities can cause side effects such as exploitation in crime, unauthorized accidental disclosure, and unfair social stigma. This possibility has a negative effect on people's behavior and thinking and can be a significant loss.

4.3 Filter bubble effect

Personalization of the platform or website, such as portals and apps, and the provision of personalized media content and news services based on the individual's Internet and mobile usage behaviors can affect the formation and maintenance of public forums for exchanging diverse views and ideas.

A filter bubble refers to a personal information space formed by a personalized algorithm, and its contents are filled differently according to algorithm sources such as searching according to the propensity and preference of the information subject. In the filter bubble, citizens check their viewpoints and values and are exposed only to content, so only their opinions and values are amplified. Being trapped in this filter bubble limits your chances of self-development because you receive limited information, limit your exposure to different people with different perspectives and perspectives, and limit the range in which you can communicate and exchange your thoughts and perspectives. Concerns have been raised that democracy can be undermined.

5. CONCLUSION

In this paper, we studied to find out the current state of big data used in science and technology policy. Big data is a necessary format for modern people living in the industrial information age. It can also provide information on the implementation of government policies used in terms of science and technology policy. This paper explores the concept of big data and uses it in science and technology policy. There are limitations and ways of overcoming those problems or limitations that arise while using big data.

ACKNOWLEDGEMENT

The present Research has been conducted by the Research Grant of Kwangwoon University in 2019.

REFERENCES

[1] Brutlag ,J. Speed Matters. Google. Research Blog. June 23, 2009; http://googleresearch.blogspot. \com/2009/06/speed-matters.html

[2] Gray, W. D., and Boehm-Davis, D. A. Milliseconds matter: An introduction to microstrategies and to their use in describing and predicting interactive behavior. Journal of Experimental Psychology:Applied 6, 4 (2000),322–335.

[3] Dasu, T., and Johnson, T. Exploratory Data Mining and Data Cleaning. John Wiley & Sons, Inc., New York, 2003.

[4] Raman, V., and Hellerstein, J. M. Potter’s wheel: An interactive data cleaning system. In Proceedings of the 27th International Conference on Very Large Data Bases (Rome, Sept. 11-14). Morgan Kaufmann, San Francisco, 2001, 381–390.

[5] Cypher, A. Watch What I Do: Programming by Demonstration. MIT Press, Cambridge, MA, 1993.

[6] Gulwani, S. Automating string processing in spreadsheets using input-output examples. In Proceedings of the 38th annual ACM SIGPLANSIGACT Symposium on Principles of Programming Languages (Austin,

Jan. 26-28). ACM Press, New York, 2011, 317–330.

[7] Kandel, S., Paepcke, S., Hellerstein, J. M., and Heer, J. Wrangler: Interactive visual specification of data transformation scripts. In Proceedings of the 2011Annual Conference of Human Factors in Computing Systems (Vancouver, May 7-12). ACM Press, New York, 2011, 3363–3372.

[8] Horvitz, E. Principles of mixed-initiative user interfaces. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Pittsburgh, May 15-20). ACM Press, New York, 1999, 159–166.

[9] Kang, H., Getoor, L., Shneiderman, B., Bilgic, M., and Licamele, L. Interactive entity resolution in relational data: A visual analytic tool and its evaluation. IEEE Transactions on Visualization & Computer Graphics 14, 5 (2008),999–1014.

[10] Robertson, G. G., Czerwinski, M. P., and Churchill, J. E. Visualization of mappings between schemas. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Portland, April 2-7). ACM Press, New York, 2005, 431–439.