November-December 2019
ISSN: 0193-4120 Page No. 2460 - 2465

Case Studies and Strategy Design of Bigdata Center for Medical Service based on Bigdata and Artificial Intelligence

Da Hye CHOI[1], Sangwon LEE[2], Hyung Su KIM[3]
[1]Seoul Asan Hospital, Republic of Korea, flrain2@naver.com
[2]Wonkwang University, Republic of Korea, (corresponding author) sangwonlee@wku.ac.kr
[3]Hansung University, Republic of Korea, hskim@hansung.ac.kr

Article Info
Volume 81
Page Number: 2460 - 2465
Publication Issue:
November-December 2019

Article History
Article Received: *5 March 2019*
Revised: *18 May 2019*
Accepted: *24 September 2019*
Publication: *12 December 2019*

Abstract

The Bigdata center was established to overcome its time-consuming and subjective limitations in finding problems through traditional business analysis methods such as interviews and surveys, observations and case analysis. In the end, we will select a study that will have more ripple effects than analyzing scattered data in hospitals. However, there are currently problems that need to be solved in order to utilize the nation's health care Bigdata, making it difficult to analyze the data. This study analyzes the government's orientation, overseas cases, and capabilities within hospitals to derive problems, find Insights, and present strategic directions from a mid- to long-term perspective. After identifying the difficulty level of the study, priority can be selected to help reduce research period and research cost through proper staffing and efficient support.

Keywords: *Artificial Intelligence, Bigdata, Bigdata Center, Medical Service*

1. INTRODUCTION

At the national level, the use of Bigdata [1-11] is expected to be a major driver of solving current problems and creating new values. The medical and health sectors are regarded as areas of potential value and high utilization of Bigdata. McKinsey analyzed that Bigdata generated $300 billion in annual economic benefits in the U.S. healthcare sector and $250 billion in the European public sector. In South Korea, medical Bigdata forecast a reduction of about 10.7 trillion won in government spending. However, data analysis is not easy as there are currently problems that need to be solved to utilize the nation's health care Bigdata. Health care data is collected from both the public and private sectors. Public data can be divided into genomic data, billing and administrative data, and survey data. In the private sector, medical institutions can be divided into clinical data and stream data collected during patient care. While the amount of data collected in the public and private domains is abundant, there are limitations to building and utilizing the data. This is because there is a lack of connectivity of data distributed by agencies, a lack of legal, technical, and policy mechanisms, and the creation of value through open data is also limited. To address these limitations, the Health and Medical Bigdata Policy Review Committee has been set up to discuss ways to promote Bigdata in health care.

2. CASE STUDIES OF KOREAN BIGDATA ENVIRONMENTS

2.1 Understanding the Government's Direction for Healthcare Bigdata

With the aim of improving health care quality, improving health care policies, and making safe and transparent use of health care Bigdata, five detailed initiatives will be implemented. The first of the detailed initiatives is the establishment of a platform to link health care Bigdata. The government intends to process data so that individuals cannot be identified by establishing a platform that links health care data to Bigdata, and to limit the risk of data transfer and re-identification through operation of closed environments. Second, it will prepare health and medical Bigdata system and guidelines. The government intends to prepare guidelines for requesting and providing information, assessing public nature and assessing security threats to overcome the lack of a unified data protection and data link base. The third is the promotion of health care Bigdata governance. Build health

care Bigdata governance to systematically discuss health care Bigdata principles. Fourth, the government is pushing to improve laws related to health care and Bigdata. It will push for the improvement of laws on the protection and utilization of Bigdata in health care and prepare measures to improve laws on principles, rights, protection and utilization procedures, responsibilities and punishments. Finally, the task of establishing a system to support Bigdata in health care. It is planning to form a research support system for health care Bigdata and provide R&D support. In doing so, health care Bigdata should be utilized for public purposes, and a discussion structure based on citizen participation and expertise should be established and the rights of informants should be fully protected based on current laws.

The budget for detailed projects will be divided into two areas, the information service industry and R&D, with a total budget of 10,337 million won. It will invest 46.01 million won in building platforms that are included in the information service industry, and 2425 million won in the analysis data network for training guidelines for health care Bigdata. The R&D project has a budget of 1134 million won for pilot projects, 19 million won for main projects and 3234 million won for governance support. The task of investing in pilot R&D of Bigdata in health care is to develop a stroke outcome and management model using Bigdata (200 million won), personalize the risk of chronic diseases in Korea, develop a preventive care service model (200 million won), develop a septicemia prediction model using biosignal and clinical information Bigdata (200 million won), and assess artificial [12-15] health exposure and health conditions using Bigdata.

2.2 Data Status of Health Care Public Institutions

The reality is that the basic data of medical Bigdata is not the hospital's own data, but it is inevitable to use public data. Although there are analyses through SNS, it is efficient to use public data when it comes to precision diagnosis. Public data can be used to analyze specification-based analysis, medication use analysis, medical management-related indicators generation, patient-based analysis, cohort research, and patient control study [16-25]. Medical data released by public institutions are provided by such organizations as the National Health Insurance Corporation, the Health Insurance Review and Assessment Service, the Korea Centers for Disease Control and Prevention, the National Institute of Health, the National Cancer Center and the Health Industry Promotion Agency. The National Health Insurance Corporation is in charge of managing qualifications, managing insurance premiums, insurance benefits and long-term care for the elderly. Data held by the National Health Insurance Corporation can be processed to generate cohort data and used and reprocessed in various forms. The Health Insurance Review and Assessment Service uses data on costs, salaries, manpower and facilities through primary data submitted by nursing institutions. The National Institutes of Health analyzes human genes generated from the tasks for supporting the Ministry of Health and Welfare, and implements customized medical treatment by utilizing the integrated management system for collecting and sharing information. The National Cancer Center is the only cancer patient data in Korea, providing data on cancer registration statistics, examination patient cohort data, history management and various information on cancer screening subjects by year. The Health Industry Promotion Agency provides statistics on the health industry with a variety of search services. It is providing statistical information on various publications and reports by the Korea Institute for Health Industry Promotion.

3. CASE STUDIES OF INTERNATIONAL BIGDATA ENVIRONMENTS

A literature survey was conducted on RISS, Google Scholar, the National Library of Korea, the National Assembly Library and various other websites to search for and utilize papers including the case of Bigdata in advanced medical services. Based on the study of health systems and services, we intend to utilize data on four different perspectives on value-added values, including patients, suppliers, policies, and research, based on good practice. Based on the results of the review, the assessment of value added and quality generated as a result of the study, such as quality and effectiveness of treatment, sustainability of the health system, eradication of chronic diseases and support of a healthy lifestyle were conducted.

3.1 Comet K-Project DEXHELPP - Austria

Austria aims to develop new methods, models and technologies to analyze the health care system's status, predict future developments, and compare scenarios based on various interventions in the healthcare sector. It covers a variety of routinely collected data sources from infectious diseases to cost data and is part of health system and service research. Two Austrian government departments and the Vienna City-supported Austrian COMET-K Study, DEEXHELPP focuses on questions ranging from data security and data management, statistical methods, causal relationship reasoning, mathematical decision analysis modeling and simulation to data visualization and public health, as well as various related interdisciplinary subjects.

3.2 The Shared Care Platform - Denmark

Denmark is an Internet IT platform that supports cross-sector collaboration in the medical field by facilitating coordination among working-level workers, local governments and hospitals, focusing on patients suffering from chronic diseases, and will gradually expand its coverage. Supported by a state fund for chronic diseases, the platform was developed in cooperation with IBM in southern Denmark. Data collected regularly from individual IT systems of health

care providers, corresponding to health system and service research, is collected from individual IT systems of health care providers, which are the basis of patients' general treatment plans, and patients have access to their data from computers and smartphones, and can add data to the system. Data stored on this platform can be printed out and analyzed, and healthcare resources can be used intensively in patients who are not well treated.

3.3 E-Estonia National Identity Scheme - Estonia

Estonia's goal is to reduce bureaucracy in doctors' workflow, distribute work hours more efficiently, allow doctors access to medical information and make health care more patient-friendly. To promote and develop e-solutions, the Estonia e-Health Foundation was established and various health studies such as health records, digital registration, digital images and e-prescribing were developed during the 2005-2008 period. The e-health record is a nationwide system that integrates data from multiple healthcare providers into a common patient record, consisting of information on diagnosis, contraceptive visits, examinations, inpatient treatment and prescribed medicines. With a centralized system, all hospitals and pharmacies are connected to the system and service is available based on recorded prescriptions.

3.4 ARNO Observatory - Italy

Italy aims to combine and integrate its databases into massive patient care data such as medication prescriptions, hospital vacancies, pharmacy home services, diagnostic testing and laboratory analysis. The application of the Arno observation system is related to health system and service rescan, surveillance of public health and epidemiology, which processes routinely collected clinical and patient care data. The ARNO observation system, a network system for epidemiological and economic monitoring, was implemented by CINECA, a nonprofit consortium comprising Italy's 70 universities, four research institutes and the Ministry of Education. Information from this system can be created by cross-checking databases that are linked to different data bases and coming from different areas. ARNO observation systems can also compare epidemiological and economic indicators by linking data from different sources.

3.5 HES(Hospital Episode Statistics) - The United Kingdom

HES is a patient record-based system designed as a data warehouse with detailed information on outpatient reservations and emergency room inmates who were admitted for use for secondary nonclinical purposes, including the basis for hospital admissions in 1987. Data including all acute hospitals, primary care and mental health in the UK are subject to this system. Each patient record contains clinical, patient, administrative and geographical information, and provides inpatient care records since 1989, outpatient records since 2003 and A&E data after 2007.

3.6 YODA(Yale University Open Data Access) - The United States

The United States was launched to facilitate access to participant-level clinical research data or clinical research reports in greater detail than journal publications to facilitate scientific research. It covers clinical research data at the participant level and comprehensive clinical research reports that apply to the scope of clinical research. The study is operated by a group of academic-level clinical researchers who partner with data holders such as Medtronic and Johnson & Johnson to access clinical trial program data. The clinical study data provided range from sensitized diseases to chronic diseases inside and 125 clinical trials as of 2016 were used in the YODA study.

3.7 CEPHOS-LINK - Finland, Austria, Romania, Norway, Slovenia, and Italy

Psychiatric hospital admissions and re-entry are high in EU countries, but the factors for re-entry are unclear. The goal of this study is to investigate psychopathic re-entry by comparing different types of interventions, focusing on the differences in the results of re-entry of adult psychiatric patients. Finland's National Institute for Health and Welfare was led by funding the EU's Seventh Framework program. CEPHOS-LINK, a comparative study of psychiatric hospital admissions through record connections of large-scale administrative data sets, is a study that examines psychiatric services conducted in six European countries, including Finland, Austria, Romania, Norway, Slovenia and Italy, from 2014 to 2017. The practice between factors such as patients, services and health systems for re-entry of discharged psychiatric patients is studied and the pattern of service out is compared.

3.8 SpainRDR(Rare Diseases Registries Research Network) - Spain

Spain aims to provide access to information and data on a central platform for health policy-making and clinical research on rare diseases. The study is led by the Institute for Rare Diseases, with a number of partners participating, including the health department, health ministry and the Center for Rare Diseases in all autonomous regions in Spain, with six medical associations, four research detours, pharmaceutical and biological engineering organizations and the Federation of Rare Diseases working together. Rare diseases are important to recruit patients who are fit to study disease causes, diseases, diagnoses or therapies, and there is a problem that research becomes complicated due to the difficulty of accessing existing information in different locations. A network established to address these problems and data is provided from two types of sources. The first is a patient registration created for a series of diseases or specific diseases to study patient treatment performance, and the second is a population-based registry established to carry out epidemiological surveys and social health plans. Data

contained in multiple registries can ultimately be combined into a single platform and used to improve diagnosis, prognosis, treatment and quality of life for patients with rare diseases.

4. STRATEGY DESIGN OF BIGDATA CENTER FOR MEDICAL SERVICE

4.1 SWOT Analysis

(Strengh) Based on the first successful medical information, research performance is being developed, there is a large number of surgical cases, and data is abundant. The center was established by selecting suitable areas as future research subjects while conducting research. It has extensive manpower pool in conjunction with Ulsan University Medical School and various research tasks. It has state-of-the-art medical equipment as a large hospital and is constantly investing in new medical devices.
(Weakness) We have a lot of data, but we need a process to transform it into structured data that is appropriate to analyze, and we need a planning plan for what kind of data we're going to analyze with a lot of data. A certain amount of performance is needed in the evaluation of medical institutions conducted by the government.
(Operations) Many companies want to collaborate with large amounts of data, and with the aging of the population, customer data is expected to grow further.
(Threats) There are limitations to the analysis of the Privacy Act and there is a lack of manpower to perform data analysis.

4.2 Research Diagnosis

In the national project and two major research parts, the national project consists of tasks that can generate synergy in cooperation with the state, and in the main study, it is likely that research can be carried out by organizing the research into tasks that can produce outstanding results by utilizing the strengths of Asan Medical Center. First of all, the government-run research task is to carry out research that can be developed in collaboration with external companies, such as joint research on convergence of industries, universities, research on future fusion medical devices, development of technology for early diagnosis of pancreatic cancer, development of technology and customized treatment strategies, establishment of clinical linked medical devices development centers, provision of clinical linked platforms, development industries, nonclinical development centers, and artificial intelligence medical imaging businesses. The main research focuses on the research that can be analyzed with data inside hospitals. It aims to conduct cancer research that analyzes factors that contribute to various cancer outbreaks, to develop prevention and treatment through neurological diseases and comprehensive research, to identify the brain nerves before developing symptoms of obesity, diabetes and arteriosclerosis through metabolic diseases, to analyze the root causes and results of incurable diseases such as cancer, brain disease and heart failure through genetic research, and to develop stem cells using research. It is required to expand the scope of Bigdata analysis based on the center organization research that is being strategically taken by transplant sectors, endoscopes, and Asan Medical Center, and establish mid- to long-term strategies for research according to the degree of collaboration and value creation.

4.3 Propose Bigdata Strategy

The Bigdata Center at Asan Medical Center in Seoul aims to promote the provision of converged medical services and technological development in line with changes in the medical paradigm, overcoming the limitations of traditional approaches. It aims to establish a medical Bigdata-oriented neoclassical system with a vision of "No trials but no failures." Data science and AI technology with high general purpose will converge with the medical field to create new service value. In the short term, the government hopes to achieve results through internal cooperation, focusing on data held by Asan Medical Center, and expand the volume through links with other medical institutions when bills on Bigdata are drawn up at a national level in the future. It will establish a research infrastructure that integrates long-term distributed human and material resources related to Bigdata and a data infrastructure for utilizing structured and unstructured large-capacity clinical data. It will provide consulting support for policies and licensing procedures on various legal and ethical issues arising from research and commercialization using Bigdata. In order to commercialize Bigdata R&D, it will encourage idea discovery and research advice and help match research and business entities such as floor researchers and businesses to establish a medical Bigdata ecosystem that leads to R&D, and commercialization, investment and research and development. It also implements an integrated research resource management system and supports IT research resource management related to optimization, maintenance, monitoring, and operation of the established management system. The medical institution carries out the national regulatory agency policy tasks for the acquisition and management of personal information related status surveys and certification, and the establishment and improvement of national policies [26-27].

5. CONCLUSION

It is difficult to find research and decide how to lead Bigdata analysis. In order to review the possibility of conducting research, it is important to prepare evaluation criteria considering the situation at Asan Medical Center in Seoul. For example, the Ministry of Government Administration and Home Affairs prepares an evaluation criteria table to select a group of candidates when evaluating candidates for public. The assessment criteria table reviews the need for implementation, ripple effects, urgency of implementation, feasibility, data collection, and model scalability, and evaluates the feasibility of implementation by item. It is necessary to establish a baseline in finding analysis tasks and

November-December 2019
ISSN: 0193-4120 Page No. 2460 - 2465

evaluating task candidates through these pre-preparatory processes. They also need to think about how Bigdata analytics can help them solve their problems. Asan Medical Center in Seoul is a general hospital in the country's Big Five with abundant data and diverse manpower. Depending on the direction in which they are used, they will be able to create a great synergy effect.

ACKNOWLEDGEMENT

This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIP). (No.2018-0-00705, Algorithm Design and Software Modeling for Judge Fake News based on Artificial Intelligence)

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2018S1A5A2A03038738, Algorithm Design & Software Architecture Modeling to Judge Fake News based on Artificial Intelligence)

REFERENCES

1. S. Park, J.S. Hwang, and S. Lee, **A Study on the Link Server Development Using B-Tree Structure in the Bigdata Environment**, Journal of Internet Computing and Services, Vol. 16. No. 1. pp. 75-82, 2015.
2. S.B. Park, S. Lee, S.W. Chae, and H. Zo, **An Empirical Study of the Factors Influencing the Task Performances of SaaS Users**, Asia Pacific Journal of Information Systems, Vol. 25. No. 2. pp. 265-288, 2015.
3. S. Park, and S. Lee, **Big Data-oriented Analysis on Issues of the Hyper-connected Society**, The E-Business Studies, Vol. 16. No. 5. pp. 3-18, 2015.
4. Jumin Lee, S.B. Park, and S. Lee, **Are Negative Online Consumer Reviews Always Bad? A Two-Sided Message Perspective**, Asia Pacific Journal of Information Systems, Vol. 25. No. 4. pp. 784-804, 2015.
5. J.K. Kim, S.W. Lee, and D.O. Choi, **Relevance Analysis Online Advertisement and e-Commerce Sales**, Journal of the Korea Entertainment Industry Association, Vol. 10. No. 2. pp. 27-35, 2016.
6. S.W. Lee, and S.H. Kim, **Finding Industries for Bigdata Usage on the Basis of AHP**, Journal of Digital Convergence, Vol. 14. No. 7. pp. 21-27, 2016.
7. S. Lee, and S.Y. Shin, **Design of Health Warning Model on the Basis of CRM by use of Health Big Data**, Journal of the Korea Institute of Information and Communication Engineering, Vol. 20. No. 4. pp. 1460-1465, 2016.
8. M. Nam, and S. Lee, **Bigdata as a Solution to Shrinking the Shadow Economy**, The E-Business Studies, Vol. 17. No. 5. pp. 107-116, 2016.
9. S.H. Kim, S. Chang, and S.W. Lee, **Consumer Trend Platform Development for Combination Analysis of Structured and Unstructured Big Data**, Journal of Digital Convergence, Vol. 15. No. 6. pp. 133-143, 2017.
10. Y. Kang, S. Kim, J. Kim, and S. Lee, **Examining the Impact of Weather Factors on Yield Industry Vitalization on Bigdata Foundation Technique**, Journal of the Korea Entertainment Industry Association, Vol. 11. No. 4. pp. 329-340, 2017.
11. S. Kim, H. Hwang, J. Lee, J. Choi, J. Kang, and S. Lee, **Design of Prevention Method Against Infectious Diseases based on Mobile Bigdata and Rule to Select Subjects Using Artificial Intelligence Concept**, International Journal of Engineering and Technology, Vol. 7. No. 3. pp. 174-178, 2018.
12. I. Jung, H. Sun, J. Kang, C.H. Lee, and S. Lee, **Bigdata Analysis Model for MRO Business Using Artificial Intelligence System Concept**, International Journal of Engineering and Technology, Vol. 7. No. 3. pp. 134-138, 2018.
13. S. Kim, S. Park, J. Kang, and S. Lee, **The Model of Bigdata Analysis for MICE Using IoT (Beacon) and Artificial Intelligence Service (Recommendation, Interest, and Movement)**, International Journal of Engineering and Technology, Vol. 7. No. 3. pp. 314-318, 2018.
14. S.H. Kim, J.K. Choi, J.S. Kim, A.R. Jang, J.H. Lee, K.J. Cha, and S.W. Lee, **Animal Infectious Diseases Prevention through Bigdata and Deep Learning**, Journal of Intelligence and Information Systems, Vol. 24. No. 4. pp. 137-154, 2018.
15. S. Lee, and I. Jung, **Development of a Platform Using Big Data-Based Artificial Intelligence to Predict New Demand of Shipbuilding**, The Journal of The Institute of Internet, Broadcasting and Communication, Vol. 19. No. 1. pp. 171-178, 2019.
16. H. Hwang, S. Lee, S. Kim, and S. Lee, **Building an Analytical Platform of Bigdata for Quality Inspection in the Dairy Industry: A Machine Learning Approach**, Journal of Intelligence and Information Systems, Vol. 24. No. 1. pp. 125-140, 2018.
17. Y. Shon, J. Park, J. Kang, and S. Lee, **Design of Link Evaluation Method to Improve Reliability based on Linked Open Bigdata and Natural Language Processing**, International Journal of Engineering and Technology, Vol. 7. No. 3. pp. 168-173, 2018.
18. T. Minami and K. Baba, **A Study on Finding Potential Group of Patrons from Library's Loan Records**, International Journal of Advanced Smart Convergence, Vol. 2, No. 2, pp. 23-26, 2013.
19. S.H. Kim, M.S. Kang, and Y.G. Jung, **Big Data Analysis using Python in Agriculture Forestry and Fisheries**, International Journal of Advanced Smart Convergence, Vol. 5. No. 1, pp. 47-50, 2016.
20. W.Y. Kim, **A Practical Study on Data Analysis Framework for Teaching 3D Printing in Elementary**

November-December 2019
ISSN: 0193-4120 Page No. 2460 - 2465

School, International Journal of Internet, Broadcasting and Communication, Vol. 8, No. 1, pp. 73-82, 2016.

21. H.C. Kang, K.B. Kang, H.K. Ahn, S.H. Lee, T.H. Ahn, and J.W. Jwa, **The Smart EV Charging System based on the big data analysis of the Power Consumption Patterns**, Vol. 9, No. 2, pp. 1-10, 2017.
22. Y.I. Kim, S.S. Yang, S.S. Lee, S.C. Park, **Design and Implementation of Mobile CRM Utilizing Big Data Analysis Techniques**, The Journal of The Institute of Internet, Broadcasting and Communication, Vol. 14, No. 6, pp. 289-294, 2014.
23. S.J. Oh, **Design of a Smart Application using Big Data**, The Journal of The Institute of Internet, Broadcasting and Communication, Vol. 15, No. 6, pp. 17-24, 2015.
24. J. Kang and S. Lee, **Algorithm Design to Judge Fake News based on Bigdata and Artificial Intelligence**, International Journal of Internet, Broadcasting and Communication, Vol. 11, No. 2, pp. 50-58, 2019.
25. J. Kang and S. Lee, **Strategy Design to Protect Personal Information on Fake News based on Bigdata and Artificial Intelligence**, International Journal of Internet, Broadcasting and Communication, Vol. 11, No. 2, pp. 59-66, 2019.
26. Hussain, H.I., Kamarudin, F., Thaker, H.M.T. & Salem, M.A. (2019) **Artificial Neural Network to Model Managerial Timing Decision: Non-Linear Evidence of Deviation from Target Leverage**, International Journal of Computational Intelligence Systems, 12 (2), 1282-1294.
27. Jabarullah, N. H., Jermsittiparsert, K., Melnikov, P. A., Maseleno, A., Hosseinian, A., & Vessally, E. (2019). **Methods for the direct synthesis of thioesters from aldehydes: a focus review**. Journal of Sulfur Chemistry, https://doi.org/10.1080/17415993.2019.1658764.