International Journal of Engineering & Technology, 8 (1.4) (2019) 395-402



International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET



Research paper

Suggestions for Digital Archives of Korean Christianity

John So¹*, Byoung-Bae Hwang²

¹Myongji University
² Hyupsung University
*Corresponding author E-mail: soyohan@yonsei.ac.kr

Abstract

The purpose of this study is to assist researchers by making it quicker and easier to access data related to Korean Protestant history. The findings can be applied broadly to the field of Protestant history and will help researchers digitalize historical materials and study them systematically. The most important problem addressed in this study, in addition to the issue of storing and accessing data electronically, is the digitalization of historical data and the way in which electronic documents in humanities fields can reflect various types of information and data in a single document. We should classify the content of Korean Protestant historical data with this in mind. This study aims to create a fast and accurate research environment suited to the IT age through electronic archives that are well maintained and accessible to anyone. This paper also discusses ways of labelling primary sources to clearly differentiate data—in particular, by tagging search words in related major themes. Archives need to focus on mutual security, not one-sided support of researchers.

Keywords: Korean Christianity, primary source, interdisciplinary integration, GPS tagging, digital archives

1. Introduction

Korean Christianity is a Western religion that was introduced during modern times. Western Protestant missionaries entered Korea and undertook various activities. They reported on these activities through mission reports, diaries, newspapers, and magazines. These materials provide important information that is relevant not only to the history of the Protestant church, but also to the modern and contemporary history of Korea.

Around 130 years ago, the missionaries' data were lost due to various archiving failures and threats, including the Korean War. Even when the data were gathered and organized, there were many restrictions on access. These constraints have contributed to a decline in research studies and intellectual breakthroughs.

The purpose of the present study is to aid research by making it quicker and easier to access data related to Korean Protestant history and associated fields in the humanities. The findings can be applied more broadly to the field of Protestant history, to help researchers digitalize historical materials and study them systematically.

In other research areas, work is currently underway to convert humanities data into electronic documents and research results. For some time, researchers have also investigated ways to preserve electronic records [1]. Other studies have aimed to enhance complementary systems by using Office Open XML and encrypting data to efficiently manage electronic records. This study focuses on security, rather than on accessing electronic documents [2].

The most important problem addressed in this study, in addition to the issue of storing and accessing data electronically, is the digitalization of historical data and the way in which electronic documents in humanities fields can reflect various types of information and data in a single document [3].

This paper examines the current use of and ways of using Protestant historical data, aiming to provide concrete alternatives.

2. The Classification of Korean Christianity Data

Charles Dollar lists eight conditions for determining the value of electronic records. The first calls for "read completion" and readability. The second states that an electronic record should be able to express information symbolized on a monitor or in printed form. Third, it must be possible to discriminate and distinguish between records. For example, information entities must be identifiable using attributes unique to their organization, classification, and other categories. Fourth, electronic records should be composites of all types of information, including consistent logic, metadata, and links. Fifth, all electronic devices should be searchable. The sixth condition is awareness; electronic records must be recognized by both PCs and people. Seventh, the electronic record itself must be viewable offline, using devices such as monitors. Finally, the eighth condition is that the record must be able to be stably stored for a long period of time. [4] One of Dollar's suggestions is that electronic records must contain a diverse range of information; this is now becoming an issue. Given these conditions, how should we classify the content of Korean Protestant historical data, the focus of this study?



2.1. Proposal for classifying Korean Christianity data

As encyclopedias became digitized, existing paper encyclopedias became less useful. For example, a search function produces information faster and more accurately than a manual search through a paper book. The main reason that digital encyclopedias have replaced existing paper media and succeeded as electronic documents is that the primary search terms and content are directly defined and linked to data retrieval. Historical data are extremely limited, in that they provide definitions and simple information in the form of words, as in the case of encyclopedias. Historical data should therefore be covered by various related search functions beyond simple classification and organization.

What is the best way to distinguish between historical data? When we look at the Korean History Elements Integration System (www.kinds.or.kr), information is generally classified into people, time, terms, spaces, events, and literature. Searches are often based on direct search terms. However, knowledge-related information plays an important role when real searches are needed. For example, when investigating a specific person or space, it may not be necessary to search, as unnecessary information can be bypassed. If an author works directly on the data and confirms it, he or she will obtain various types of data; search results may not be very useful. For example, when searching for historical records on "church construction," only articles with the terms "church" and "architecture" will be recovered; words linked to church building or the church itself will not searched for. Information that can identify the context is not searchable by itself. There is thus a need to develop a system that integrates and links historical elements and content. To do this, systems must assist with retrieval and provide researchers and users with feedback on data retrieval; existing archives merely provide one-sided data.

To achieve this, JSTOR recently announced a service called Text Analyzer Beta (https://www.jstor.org/analyze/). This service provides a tool to search for documents already in the user's possession. When users upload individual documents (such as PDFs, text documents, or images), they can be searched and matched with materials owned by JSTOR. This is clearly a useful tool for researchers; it not only extends beyond traditional search methods, but also expands the scope of research. Although this may be helpful for individual researchers, students who want to study the materials or users looking for related topics may have to go through the same search process or limit their searches.

Figure 1 illustrates how the JSTOR search function should be approached. First, one drags and drops the PDF, text, or image file and makes it easy to access by uploading it.



Figure 1: Interaction between electronic archives and readers

As an example, Figure 2 presents a primary source covering martyrdoms in Korean church history. The material was recorded in English and stored as a PDF file. Both the recognition rate of this medium and the number of studies that use it have been tested. Figure 2 shows a document that was uploaded using JSTOR's Text Analyzer.

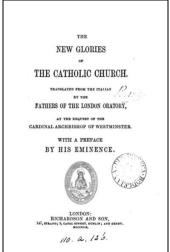


Figure 2: Samples of martyrdom material in Korean church history

The above data involve Korean church history and the martyrdom of Catholics. The following analysis was carried out.

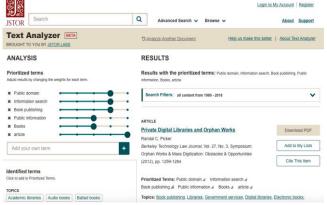


Figure 3: Analysis of Korean Catholic martyrs: usage data

Figure 3 shows the research papers and patterns of use. This guideline allows the research trends of related studies to be examined and facilitates analyses of searches. The tool allows researchers to analyze previous research, thus enriching the research.

2.2. Proposal for using materials on Korean Christianity

Previously, we examined classification methods, data retrieval, and ways of analyzing historical records. This tool, when sorting or searching for historical data, searches all of the data, rather than just recognizing the keywords. In this way, the limitations of word searching can be resolved. More specialized searches, involving Christian, Korean, and historical data, can be made more precise through specialized search functions. Figure 4 shows that archives provide an opportunity to get help, rather than just delivering data to researchers and users unilaterally. The advantage of this system is that it allows users to search data in this area more precisely and integrally, using search tags added by other experts.

If researchers are encouraged to use search tools to upload their personal collections of data into electronic archives and to share that data, they will be able to help others collect data.

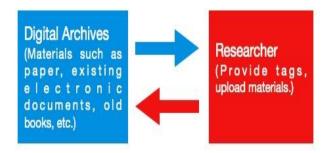


Figure 4: Interaction diagram of electronic archives and readers

Figure 4 illustrates a good alternative approach that adds various search terms to search engines for historical data. During the process of storing historical records, researchers must add marks (recognition) and specific symbols to specific terms, such as place names; this is called a markup language [5]. MARKUS (http://chinese-empires.eu) has a project called "Communication and Empire: Chinese Empires from a Comparative Perspective," which searches for ancient Chinese documents. This website works on PC, MAC, or other operating systems. It even works on individual devices. It aims to classify people, places, and specific words, thus enabling researchers to approach and interpret ancient Chinese documents. The researchers' data, including electronic tags, can be uploaded separately.

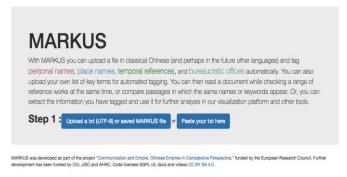


Figure 5: Interaction diagram of electronic archives and readers

This project, sponsored by the University of California, San Diego Academic Senate, will take place at Harvard University, National Taiwan University, and the University of Birmingham. The program initially aims to convert all of China's ancient documents and various historical records in Chinese into text files. The homepage has its own Hanja recognition program; researchers can add additional security features if they encounter a lack of related materials or searches. In addition, it continuously corrects bugs and errors in the Chinese character recognition program and increases the recognition rate. In this way, it can help to develop a collection, arrangement, and classification system for Christian history data in the future.

However, it would be ideal if this project could be extended to include the history of Korean Christianity, which is a specific field. Many tasks need to be completed to develop its own program and collect a vast amount of data. It will be very advantageous if Korean Christian historical materials can be accessed alongside the collected and analyzed data on the MARKUS homepage. This model will help researchers identify, access, and use a wide range of historical data from the large data collections of small laboratories and institutions. As Figure 6 illustrates, it is possible to browse the list of materials on the MARKUS homepage and help ensure that this open data become available to anyone.



Figure 6: Primary Sources of MARKUS

Up to this point, we have explored ways of analyzing the historical data in academic papers using samples from other institutes. The first challenge is to create text versions of the historical data in papers. In both English and Korean, these tools can translate information into text in the same way so that data held by an individual can be transformed through the OCR PDF process. Another approach is to borrow a text tool from MARKUS or JSTOR. These tools are patented in English and Chinese, allowing easy access and use for text work. Second, it is necessary to strengthen the search function. Strengthening the search function requires analyzing the primary data of related researchers or tagging specific words by grasping the overall character of the document. These tags can be added to the PDF search function; alternatively, the data can be uploaded and sorted by building a homepage system that inserts tags, using the MARKUS method mentioned above.

In the next section, we will explore the whole structure of Korean Christianity data, explaining why it is important to collect such historical data and how to apply it in real life.

3. The Importance of Primary Sources Related to Korean Christianity

3.1. Research Trends in Korean Christianity

Early Korean Christianity was recorded by Western missionaries in their mission reports. One representative early church work is Claude-Charles Dallet's Catholic Church of Korea. Protestant denominations also have missionary reports and have organized data and recorded various events [6]. This field of study is generally divided into two parts. One part involves the study of missionaries and other overseas scholars, while the other focuses on domestic scholars. A number of overseas studies have been based on primary sources and existing missionary data; these are included within the area of Korean studies [7]. Studies of overseas scholars and missionaries include the Lecture Book by Dr. G. H. Jones (Drew, 1910) and the Joseon Christian Brief History by Dr. Harry A. Rhodes [8]. Examples of domestic Korean studies include Joseon Christian and Diplomatic History by Neung-Hwa Lee (written from a political diplomacy perspective) and Dr. George Paik's The History of Protestant Missions in Korea 1832–1910 (written from the perspective of missionary history) [9]. Building on this research, second-generation domestic studies have looked at church history from a Korean perspective [10]. The Korean viewpoint emphasizes subjectivity, reflected in the Korean people's active and voluntary acceptance; this argument originated in "the internal development theory" proposed by Japanese colonial scholar Fukuda Tokuzo, who claimed that the Joseon economy was stagnant and that Japan had taken the lead in modernizing the Joseon Dynasty. This theory argued that acceptance was inherent among the Korean people, who voluntarily modernized the economy. Scholars influenced by this argument included Lee Kwang-rin and Lee Man-yul, who specialized in Korean history. Their studies focused on the potential of Koreans. However, it is true that their theory was based on a refutation of colonial history. In addition, Kyung-Bae Min viewed Korean church history from the viewpoint of internal faith experiences and external social participation, centered on the Great Revival movement of 1907; this viewpoint made it difficult to include the whole of Korean church history [11]. Jae-Yong Ju has studied popular ministerial chiefs, but this topic demands further study. Spencer J. Palmer, Robert E. Buswell, Timothy S. Lee, and David Chung are all overseas researchers. Although they are introducing Korean church history in English, there is no difference between their research and church history studies in Korea [12]. Recently, however, research on Korean church history has been revived, with new theories and arguments emerging as papers rather than books.

3.2. The importance of primary source research

Based on previous studies, various research methods have been used to research Korean church history, and these methods have produced results. However, in order to use this data to contribute to the field, researchers must identify major errors in the existing research; they need a system for verifying data [13]. As the image below illustrates, the content of previous Korean church histories has been inaccurately cited, due to mistakes in primary source citations. Figure 7 presents the process by which Koreans became Protestants for the first time.

11.1. Anonymous Christians Converted in China

An interesting fact is found about the first conversion of a Korean to the Protestant faith. According to the letter written by Dr. Joseph Edkins dated 15th April 1864, who dwelled in Shanghai in 1864, when Dr. Dugeon arrived the church session examined those applicants for admission to the Church. Four persons were selected to be baptized, among them three persons were those who were serving at the official residence of the Royal Prince of Korea who sojourned in Shanghai. Since no detailed information as to the persons is found in the letter, it is difficult to judge whether or not they were Koreans. But because it was the old custom that the Royal family of Korea who sojourned in China always took along attendants with them, it is most probable that those three persons were Korean. Furthermore, as we have observed, there had been frequent contacts with the missionaries with the Korean annual envoys and gentlemen, it is certain that they were Korean.

If this is proven to be a fact, the history of the first baptism of the Korean Protestant Church is traced back to a day in April 1864. It was twenty years before the first American Missionary entered into Korea, and two and a half years before Robert J. Thomas was martyred. But their names were left unidentified, and no significant connection with the subsequent history was detectable. This may be a story of romantic prelude, but the presence of those three persons remained an island on the vast sea of the Church history of Korea.

 Dr. Edkins was a close friend of R. J, Thomas. A Register of Missionaries and Deputation from 1786 to 1877, 1877, List No. 492. London Missionary Society Archives.

Figure 7: Primary source misquoted in Korean church history

A close look at the source above shows that it uses information from *The London Missionary Society*. However, in the actual text of *The London Missionary Society* Vol. 28, 1864, 226, presented in Figure 8, Protestantism is described as Manchu, not Korean.

"Peking, April 15, 1864.

"MY DEAR BROTHER,—Since I last wrote, Dr. and Mrs. Dudgeon have arrived here, and Dr. Lockhart has left us. You have heard that Dr. Dudgeon passed the winter at Chefoo. He reached this place with his wife and infant in comfort and safety, on March 29th, and was thus able to spend a week here with his zealous and laborious predecessor, Dr. L., who bade us farewell on the 5th of April, and on whom we pray that Divine protection may attend on his journey homeward. He has done a good work here, as is recognised in a marked manner by his missionary brethren, and the whole of the resident foreign community.

"At a Church Meeting, held on the day of Dr. D.'s arrival, and but two hours before, several inquirers, desirous of baptism, were examined, and four

Figure 8: The London Missionary Society Vol. 28, 1864, 226.

Figure 8 illustrates the need to collect and analyze primary data. Certain titles and content, such as data related to the "first converts" associated with these historical events, must be identified using primary sources. If researchers can search for content using these main data sources, the findings will also have value in related fields. In the next chapter, I will focus on the uses of (and alternatives to) primary sources.

4. Suggestions for accessing and using primary sources

4.1. Suggestions for using primary sources

At first, primary sources were preserved in their physical form. Currently, however, attempts are being made to transfer physical sources into a database. Modern scholars have tried to improve the quality of research by storing databases in the cloud, a form of hard storage. Mass storage technology is needed to preserve selected Korean Christian historical materials. The cloud is not a form of hard storage, but a technology that professionally backs up data. However, since storage capacity is not infinitely expandable, it will be necessary to select the essential data first. What criteria will be used to select materials? First, according to CEDARS guidelines, institutional responsibility for conservation includes the following points: 1) The institution must be responsible for conservation. 2) The subject area in which the institution is interested should be clarified. 3) When creating digital data, priority is given to paper materials, which are often used. [14] The Preservation and Accessing Networked Documentary Resource of Australia provides more detailed guidelines for selecting data to be preserved. The criteria are as follows: 1) Documents written by Australians whose content is important to society, politics, and religion, and relevant to Australia. 2) When there is double data, if one set of data is published and preserved in a printed medium, it should

be saved only if electronic data are needed. If print media or microfilm data do not exist, preserve electronic data. 3) Authority and research value: Authoritative publications obtained through long-term research are given priority and shall be preserved first. These precedents suggest the right criteria to use when choosing which Korean Christian historical data to preserve electronically. The criteria proposed in this study are as follows: 1) Materials written by Korean Christians. The content should relate to politics, society, and religion. 2) Authority and research value: data and authoritative documents that have been studied over a long period of time should be converted into electronic documents. These two criteria are common data selection criteria used in other studies. So how should these materials be handled? The section discussing this issue will apply practical methods through which individuals can categorize and access sources in the following way. In addition to the data above, there must be an additional record, known as an authorized record. In this record, information will be collected to ensure that standardized forms are used for characters, events, signatures, subject names, summary names, and place names. These should be documented along with the selection of the materials mentioned above.

4.2. Suggestions for accessing primary sources

People access data using personal devices instead of carrying around heavy physical primary sources. However, data that are simply transferred into a database have the following limitations. If too much electronic data is collected and cannot be managed, it becomes difficult to access. Moreover, there may be no backup; in such cases, data that have been collected and stored in one place may be deleted all at once. Sudden problems can occur, including personal device malfunctions and access limitations due to Internet flaws. In order to preserve data and make them accessible, we must develop a system for classifying and organizing data and storing, identifying, and classifying unstructured data (characteristic of primary source users). This kind of big data can be developed as a technology and used not only in the field of church history, but in various other theological fields as well. For this reason, the basic task that theologians and scholars should undertake is to collect electronic data and refine it.

Results should be based on the primary sources categorized by the database. For example, using the OCR (text recognition) feature of a PDF file is the easiest way to create categories. The institute should also install a church history archive on its server so that many people can access the materials. Institutes should back up their primary sources for safety. In addition, if the Internet connection suddenly fails, researchers can synchronize frequently used databases with their personal devices, such as smartphones or laptops, to ensure continuity of research.

Where should these systems be stored? As previously mentioned, the cloud is the best choice for storage. The advantage of the cloud is that it is easy to build—in addition, data usage analyses and capacity expansion are both relatively easy to carry out. The governments of the US, EU, and Japan actively use the cloud because it can be easily accessed using personal cellphones. The more concrete advantages of the cloud are as follows: 1) It reduces the cost of purchasing expensive server equipment or fixing installation, operation, and backup security systems. 2) It is economical when allocating necessary equipment and time periods; it operates freely, with users paying only for what they use. 3) It is advantageous for big data analyses that require a lot of temporary resources in a data-intensive environment. 4) There is no additional software or installation burden. 5) It is possible to access the cloud using low-end mobile equipment; it is also possible to use large servers and a lot of software. 6) It is possible to share, edit, store, and synchronize documents anytime and anywhere [15]

It is impossible to study history without considering time, space, and background. In this section, we will look at how PDFs, JPGs, and other materials are listed using GPS information on an electronic map or other virtual space. For example, if a researcher or tourist is at or near a location where a historical event occurred, the electronic map will present historical data. When Adobe Acrobat 9 was released, geospatial PDF technology emerged. With this tool, researchers can add GPS information to primary sources or existing PDF articles, electronic books, and photographs; they can then be added to electronic maps. This is a very simple task. Primary source materials and data from researchers and historians can provide inspiration.

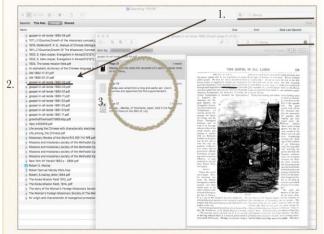


Figure 9: Examples of applied big data baseline work and primary source studies using the keywords "Maclay (Early Methodist missionary)"

- 1. The search term "Maclay."
- 2. Primary sources searched by OCR.
- 3. The sample that executed the retrieved primary source.

5. GPS tagging in primary data and thesis introductions

5.1. The application of GPS tagging to electronic materials

Because primary sources and documents are used only by interested researchers and a limited number of people, there are problems to resolve in making these data widely available. However, if augmented reality features are applied to documents, more people will have easy access to primary data and existing documents. For example, Figure 10 is a map of Jeong-dong, in Seoul. Many people rely on the combination of maps and data to navigate and find food-related and cultural destinations. It is now possible to make better use of related documents by allowing maps to be searched for data and document media. In Figure 10, the pinpoints refer to papers related to this location. When you click on a pin point, you will see papers related to the area, buildings, people, and events. These functions link the cloud homepage with journals, bringing up recorded documents. The task that remains is to systematically organize the data and articles on the map.



Figure 10: A map that applies GPS to primary sources or theses

These papers can be searched for on the map because the documents contain GPS information. In the image below, GPS information appears in the document or image; the position of the location where the work was carried out is displayed.

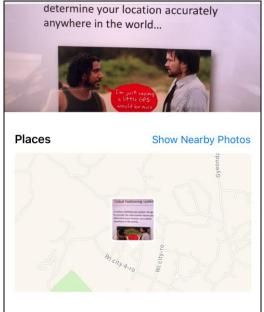


Figure 11: Data with GPS information

Although Adobe has a feature for inputting GPS data into an image or document, it is also possible to access Geosetter at http://www.geosetter.de/en/download-en/ and add GPS functionality to a file. Such files can be saved onto maps or archives and displayed on a map. Then, in addition to searching for primary data, it is possible to search for past, present, and changing processes on the map, as long as you have GPS tagging in your old photos. For example, to compare photographs taken by missionaries with modern and contemporary spaces and images, it is important to reconstruct the spaces of the past by extracting image data from the primary sources and adding GPS tags. Figure 12 is a photograph from 1899. The image shows houses owned by Western officials that cannot be seen now. You can compare the past to the present using an archive map.



Figure 12: Seoul in 1899

When using archives, the only necessary step is to implement a management phase. As it is important to ensure that initiatives are being carried out, there must be steps for managing and supervising those initiatives. In addition, errors, omissions, redundancies, and similar issues should be checked during the data search. Good management provides quick and accurate reference data in response to requests, thus building up the system.

6. Conclusion

In the humanities and history fields, the emphasis is on primary sources. For this reason, the present study focused only on collecting and classifying such materials. Due to the limitations of movement, sharing, and access, primary sources are limited to specific places and times, restricting research. In the present age, research is improved through the scientific analysis of data. This study was a practical attempt to apply these principles to Korean church history research. In particular, the combination of human research and major sources of data can improve the quality and accuracy of historical research. In particular, primary source citations in Korean church history studies may contain mistakes, which can seriously undermine scholarship and even cause misidentifications. To solve this problem, an attempt has been made to combine these different approaches. This is one of the most important human sciences as well as a related subject. It is the first step in developing historical research. It is necessary to study methodological and practical approaches, such as primary source scanning, and the PDF and OCR models (character recognition). This study aimed to create a fast and accurate research environment suitable for the IT age through electronic archives that are maintained and made accessible to anyone. This paper has also discussed ways of labelling primary sources to clearly differentiate data—in particular, by tagging search words in related major themes. Archives need to focus on mutual security, not on one-sided support of researchers.

Theologians studying church history can quickly and conveniently access the first set of data from a time period using systematic data, which allows them to grasp the era more precisely. Furthermore, if one inputs specific keywords, one will be able to distinguish by time, so as to quickly grasp the changes of the times and broaden the scope of the research. To do this, we will work on analyzing the disadvantages and merits of the real world by constructing primary data.

References

- [1] Kim M-O & Lee S-Y, "A Study on the Functional Elements for Long-term Preservation of Electronic Records," *Journal of Korean Society of Archives and Records Management*, Vol. 10, No. 2 (2010), 103–105.
- [2] Lee Y-G, Kim H-C, Jung T-Y, & Jun M-S, "Design of E-Document Management System Using Dynamic Group Key based on OOXML," *The Journal of the Korean Institute of Communication Science*, Vol. 32, No. 12 (2010), 1408.
- [3] Kim B, "A Compilation Method of Electronic Documents for Historical Records—Focused on Historical Elements" *Yeol-sang Journal of Classical Studies* Vol. 52 (2016), 215–216.
- [4] Dollar CM, Authentic Electronic Records: Strategies for Long-Term Access, Cohasset Associates Inc. (2002), 127.
- [5] Kim B, "A Compilation Method of Electronic Documents for Historical Records—Focused on Historical Elements, 205.
- [6] Dallet C-C, Histoire de l'Eglise de Corée, Paris, (1874).
- [7] Chung D, The Religious Context of Christian Beginnings in Korea, State University of New York Press, (2001).
- [8] G. H. Jones, *Lecture Book*, Drew (1910).
- [9] Paik G, The History of Protestant Mission in Korea 1832–1910, Seoul (1995).
- [10] Underwood HG, The Call of Korea, Fleming H. Revell Co., (1908); Rhodes HA, A Brief of Korean Christianity, Chosen sotokufu gakufu, (1913).
- [11] Min KB, A History of Christian Churches in Korea, Yonsei University Press, (2005).
- [12] Lee N-H, Christian Political Diplomatic History of Joseon, Keijo, (1933).
- [13] So J & Hwang B-B, "Application of Big data for Development of Korean Christian Research," *Journal of Statistical Computing and Algorithm*, Vol. 2, No. 1, (2018), 10.
- [14] Lee CS & Kim MJ, "Long-Term Preservation of Electronic Resources," Journal of Social Science, No. 13, (2001), 136–137.
- [15] Lee S, Choi J, Kang S-W, Lee K-J, Han M-H, & Kim J, "Technology for Electronic Document Management and Virtual Storage System in Cloud Environments," *Journal of Information Technology and Architecture*, Vol. 13, (2016), 182.