

## **A New Path for the Study of the Koryŏ Dynasty: Exploring the Future of Online Historical Source Archives**

*Soochan Park\**

### **Future Society and Historical Studies: At the Onset of the Age of the Fourth Industrial Revolution**

Today, at the onset of the age of the Fourth Industrial Revolution, industries and science are rapidly changing. Because of the dramatic rate of change in these areas, other areas that have been unable to keep up at the same speed seem as though they are lagging behind. Amidst these changes, the humanities are also indirectly and directly facing transformation.

In the field of history, the advancement of science and technology is not easily recognized. Historical studies have long maintained the framework of analyzing historical sources over time to derive logical results. Therefore scholars have to continue the endless task of reviewing and organizing historical sources. Even with the development of science and technology or any other changes, this framework of historical studies will not change.

Nonetheless, historical studies is a field of study that needs to be responsive to the development of science and technology. This is because analysis of materials, mainly historical sources, is of the highest priority in historical studies. Scholars have long made constant efforts to cross-

---

\* Ph.D. candidate in Korean History at Korea University

check numerous historical sources and link them together with a common thread. Their efforts have resulted in massive research results, based on which scholars are striving to produce better achievements. Unfortunately, we have reached a point where efforts are not enough to overcome the obstacle that historians are facing. The amount of accumulated research results is much too vast for one scholar to examine and analyze comparatively, and the amount of additional historical sources that are accumulating now is also too much for historians to handle.

For research on the Koryŏ (Goryeo) dynasty (918-1392), which will be examined in this paper, it is true that there is a lot less historical sources compared to the later periods. The absolute number of historical sources that have been preserved for posterity is much smaller than those from the Chosŏn (Joseon) dynasty (1392-1910) or later. However, new historical materials related to the Koryŏ dynasty, such as Chinese sources or collection of works compiled in the Chosŏn dynasty, have been surfacing. In addition, the lack of actual historical sources has led Koryŏ dynasty scholars to emphasize the general understanding of all available sources as well as comparative analysis of the sources. In most cases a general review of the whole source is necessary to find specific details.

In the study of the Koryŏ dynasty, technological transformation has not yet taken place, unlike in other fields of Korean history. Ancient history scholars have proposed the use of big data for reviewing epigraphs, while scholars focusing on the Chosŏn dynasty are already using AI to analyze the *Daily Records of Royal Secretariat of the Chosŏn dynasty* (Sŭngjŏngwŏn ilgi) and other primary sources. Yet the only technological aspect of research on the Koryŏ dynasty is that some of the historical materials from this period have been digitized and made accessible for the public via Naver Encyclopedia (Naver Chisik Paekgwa) and the Korean History Database provided by the National Institute of Korean History (Kuksap'yŏnch'an Wiwŏnhoe).

As mentioned earlier, science and technology are advancing at a remarkable speed. There are numerous ways to utilize science and technology in the humanities, and certain areas of historical studies will also ben-

efit greatly from adopting the use of science and technology. Technologies that facilitate research by increasing access to historical materials are already being applied in practice, combining with various other technologies. This paper aims to explore how new technologies of the Fourth Industrial Revolution can be applied to historical studies, more specifically for handling historical sources from the Koryŏ dynasty. To this end, first, I will examine how historical sources were handled in the past. Then, I will look at examples from foreign countries to identify the technologies that can be immediately applied in historical studies. Lastly, I will discuss the direction scholars should take in applying science and technology to historical sources for research on the Koryŏ dynasty, and the role of researchers in this process. I hope that this paper can contribute to improving the study of the Koryŏ dynasty.

### **Research on the Koryŏ Dynasty: Online Digitization of Historical Resources and Their Use**

Understanding and organizing historical sources are essential to historical research. Today, most historical sources are available in print, and they are also available in the data format, which can be accessed via electronic devices, such as computers and tablets. As an example, let us take a look at *Koryŏsa* (The history of Koyrŏ), a major primary source in the study of the Koryŏ dynasty.

A book of photoprints of *Koryŏsa* was printed by the Japanese publisher Kokusho Kankokai in 1908. In 1948, Korea's Kukje Sinmunsa published "Sega (royal annals, 世家)" from *Koryŏsa*. In 1955, Yonsei University's Institute of East Asian Studies also published photoprints of the woodblocks of *Koryŏsa*, but some parts were missing and some pages were printed upside down. Asea Munhwasa also published photoprints of *Koryŏsa* in 1972, which were organized into three volumes with each volume containing nearly 1,000 pages.

Translations of *Koryŏsa* have been published as well. In 1966, the

Classics Research Center at the North Korean Academy of Social Sciences translated and published *Koryŏsa* (North Korean translation). Dong-A University's Classics Research Center also published *Yokju Koryŏsa* (The translated and annotated history of Koryŏ) in 1973 (old Dong-A translation), which was revised and published under the title of *Kugyŏk Koryŏsa* (Korean translation of the history of Koryŏ) in 2011 (new Dong-A translation).<sup>1</sup> These were all printed publications, and they were much larger in volume compared to the photoprint books. The North Korean translation and the old Dong-A translation were both published in 11 volumes, while the new Dong-A translation amounted to a total of 30 volumes. The latter translation, in particular, was printed on B4 size paper, which is much larger than regular books and therefore physically bigger. These Korean translations have made it easier for scholars to understand *Koryŏsa*, but they have not necessarily improved the physical access to these resources.<sup>2</sup>

The groundbreaking improvement in access to *Koryŏsa* came with the internet.<sup>3</sup> Currently (last accessed by the author in May 2019), *Koryŏsa* can be viewed online via the National Institute of Korean History's data

- 
- 1 For information on *Kugyŏk Koryŏsa*, see Shin Ŭnje, "Kugyŏk Koryŏsa-ŭi tochŏn, kŭrigo han'gye (Challenges and limits of *Kugyŏk Koryŏsa*)," *Sŏkdang nonch'ong* (Journal of Seokdang Academy) 54 (November 2012).
  - 2 *Koryŏsa* became digitized in the late 1990s and the early 2000s, owing to the development of CD-ROMs. However, these CDs were expensive, and mainly purchased and used by university libraries rather than individuals. Regarding the specifics of the development of *Koryŏsa* CDs, see Yi, Namhŭi, "Koryŏsa tiji'tŏlhwa-ŭi panghyang-gwa kwaje" (The direction and challenges of the digitization of *Koryŏsa*), *Ch'ŏnggyesahak* 18 (August 2003).
  - 3 For digitalization projects of historical documents and sources from the Koryŏ dynasty, see Pak Chinhun, "Koryŏ sidae munhŏn charyo chŏngbohwa saŏp hyŏnhwang mit iyong shilt'ae-wa hyogwajŏgin hwaryong bangan (Projects for the digitalization of historical documents from the Koryŏ dynasty, their use and effective utilization method)," *Han'guk chungsesa yŏn'gu* (The Journal of Korean Medieval History) 30 (April 2011). However, there has been a great improvement since this paper was published, analyzing the situation from about a decade ago.



Fig. 1. National Institute of Korean History's Koryŏ Dynasty Historical Sources Database

base for historical sources from the Koryŏ dynasty.<sup>4</sup> According to the National Institute of Korean History (NIKH), “For phase 1, the *kujŏmbon* (literally, phrase dot edition) of *Koryŏsa* began to be provided via the internet in 2009, linked with the original images available on the Seoul National University’s Kyujanggak Institute for Korean Studies website. People were able to browse the source, which was listed by volume, date, and theme (*sega*, *chi*, *p’yo*, and *yŏljon*) for user convenience. Phase 2 in 2010 involved the provision of the *p’yojŏmbon* (modern punctuated edition) of *Koryŏsa* that includes punctuations, index-term tags (names of people, places, books, and government post), and article titles. Finally, between 2014 and 2015, phase 3 of the project was completed to reflect the latest research efforts in the field, making the Korean translation of *Koryŏsa* available online and adding a collection of diplomatic documents

4 <http://db.history.go.kr/KOREA/>

from the Koryŏ dynasty organized into different categories.”

The Koryŏ Dynasty Historical Sources Database also provides various additional services. The view by date option shows a list of historical resources from the Koryŏ dynasty for a certain date, based on the results of studies on the chronological tables from the Koryŏ dynasty. There is also an option to view sources by index term, which allows users to access documents containing the search term. Lastly, the site also provides a collection of diplomatic documents from the Koryŏ dynasty with search options that function similarly to Microsoft Excel’s filter function.<sup>5</sup>

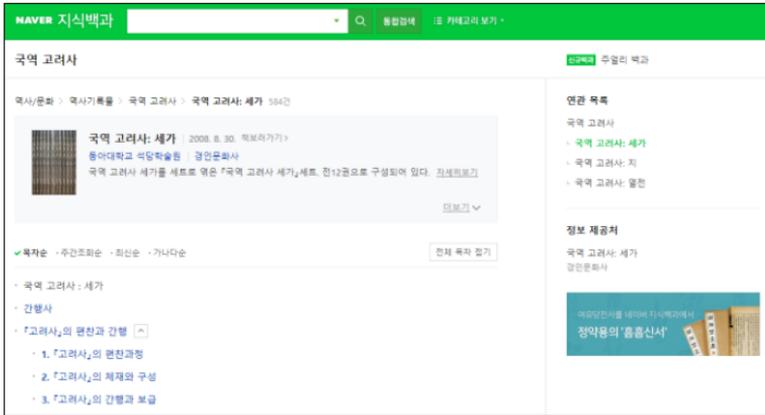


Fig. 2. *Kugyŏk Koryŏsa* on Naver Encyclopedia

Today, *Koryŏsa* is easily accessible to the public. Naver Encyclopedia began to offer *Kugyŏk Koryŏsa* (Korean translation of *Koryŏsa*) in 2012,<sup>6</sup>

5 Song Yongdŏk, 2018, “Kuksap’yŏnch’an wiwŏnhoe Koryŏ sidae saryo chŏng-pohwa hyŏnhwang-gwa kwaje” (The status and challenges of the National Institute of Korean History’s digitalization of historical sources from the Koryŏ dynasty), *Koryŏ kŏn’guk 1,100 chunyŏn kinyŏm che 53 hoe Kuksap’yŏnchan wiwŏnhoe han’guksa haksurhoeŭi kyŏnggye-rŭl nŏmŏ saeroun kilo* (The 53rd National Institute of Korean History academic conference in celebration of 1,100 years since the founding of the Koryŏ dynasty: a new path beyond the boundary), (Seoul: Kuksap’yŏnchan Wiwŏnhoe, 2018), 17-22.

6 <https://terms.naver.com/list.nhn?cid=62131&categoryId=62131>

allowing everyone with access to the Internet to view the source. Naver not only provides the photoprints but also the Korean translation of the source.



Fig. 3. Main page of the Research Institute Goryeosa (Koryŏsa Yŏn'gu) website

Korea University's Koryŏ Dynasty Research Division also provides restricted access to *Koryŏsa* via the Research Institute Goryeosa website for members of the Korean Medieval History Society (Chunsesa Yŏn'guhoe).<sup>7</sup> Although the site does not offer the Korean translation, it is useful for researchers studying the Koryŏ dynasty. For instance, it provides analyses and comparisons of various errors and important matters in *Koryŏsa*. Dates (or simply mentions of years from the sexagenary cycle) are estimated for entries without dates by examining the entries that come before and after, and contents of “Chinayŏljon” without chronological references are crosschecked with the events in “Sega (Royal Annals)” to identify when the revisions were made. The Research Institute Goryeosa website provides a revised chronology of events in the Koryŏ dynasty in addition to the original chronology to help scholars better understand the history of the Koryŏ dynasty. Researchers can view *Koryŏsa* in the chronological order as recorded in source and also in the revised chrono-

7 <http://www.khistory.org/hj/index.do>

logical order that was derived through a comparative analysis with related historical sources.

The Research Institute Goryeosa has also worked on tagging historical figures and buildings in *Koryŏsa*, linking the various words used to refer to the same person or the same building. For instance, Yang Paekyŏn (楊伯淵), whose name is also recorded as 梁伯淵 or 楊伯顏 were linked together under the same tag, while different people with the same names were separated under different tags. This project has increased the opportunities for historical sources to be analyzed in various ways.

In today's research environment, *Koryŏsa* is mainly accessed not in the book form but via PDF or other data format. As a result, it is organized using electronic devices, such as computers, rather than by hand. And *Koryŏsa* is not the only historical source that has been digitized. *Koryŏsa jŏryo* (Essentials of Koryŏ history, 高麗史節要) is also available on the National Institute of Korean History website mentioned above, in addition to collections of literary works, such as *Tonginjimun sayuk* (Collected poems of Shilla and Koryŏ literati), *Tonginjimun och'il* (Collected poems of the Koryŏ dynasty), *Chewang un'gi* (Rhyming record of emperors and kings), *Pohanjip* (Supplementary jottings in idleness), and *P'ahanjip* (Jottings to break up idleness), as well as Chinese historical sources on Koryŏ, including *Xuanhe fengshi Gaoli tujing* (Illustrated guide of Koryŏ by an envoy during the era known as Xuanhe), *Yuan Gaoli jishi* (Chronicles of the Yuan-Koryŏ relations), and *Koryŏ Parhae yumin kisa* (entries on the displaced people of Koryŏ or Parhae (Balhae) in Chinese historical sources). The Institute for the Translation of Korean Classics (ITKC, Han'guk Kojŏn Pŏnyŏkgwŏn) also provides public access to *Koryŏsa jŏryo*, *Xuanhe fengshi Gaoli tujing*, and other historical sources via the Database of Korean Classics (Han'guk Kojŏn Chonghap DB).<sup>8</sup>

The aforementioned Research Institute Goryeosa also provides access to the complete *Koryŏsa* and *Koryŏsa jŏryo*, as well as to epigraphs and inscriptions found on steles for monks (kosŭngbi) and other monuments.

---

8 <http://db.itkc.or.kr/>

In addition, *Mogŭnjip* (Collected works of Yi Saek), *Tongmunsŏn* (Anthology of Korean literature), *Kajŏngjip* (Collected works of Yi Kok), *Tongguk Yi sanggukjip* (Collected works of Minister Yi of Korea), *Ikjaejip* (Collected works of Yi Chehyŏn), *P'ahanjip*, *Tongan kŏsajip* (Collected works of Yi Sŭnghyu), *Cholgo ch'ŏnbaek* (a collection of literary works by Ch'oe Hae), and *Pohanjip* are available. The revised chronology mentioned earlier has also been applied to these historical sources as well, and having these sources organized under a common chronological system has greatly improved the ease and convenience for researchers.

These various digital services have made it easier not only for researchers but for the general public to access historical sources from the Koryŏ dynasty. Although the digitization of sources from Koryŏ has come late compared to historical sources from other periods, Koryŏ dynasty scholars are gradually applying new technologies to their sources in remarkable ways. However, the level of technological application to research is currently limited to improving access to primary sources, and therefore there is still a lot of room for technologies to be applied to other areas.

### **New Types of Digital Libraries for Historical Sources: Project Gutenberg and Chinese Text Project**

Then how should historical sources be handled in future research on the Koryŏ dynasty? There are two examples that we can use as a reference. The first is Project Gutenberg, founded in 1971 by Michael Hart, who was then a student at Illinois University.<sup>9</sup> After he received an operator's account to use the school's mainframe computer system, Hart decided to digitize literary works by turning them into e-books and distribute them for free.<sup>10</sup> The first e-book that he digitized and distributed to online

---

9 <http://www.gutenberg.org>

10 [http://www.gutenberg.org/wiki/Gutenberg:The\\_History\\_and\\_Philosophy\\_of\\_Pro](http://www.gutenberg.org/wiki/Gutenberg:The_History_and_Philosophy_of_Pro)

communities for free of charge was the United States Declaration of Independence. Since then, Project Gutenberg has continued to produce e-books of works that are in the public domain. It took over 20 years to digitize and distribute the first 100 titles, but owing to increased public interest and volunteer efforts, Project Gutenberg holds about 59,000 titles as of May 2019.<sup>11</sup>

As mentioned earlier, the titles in the Project Gutenberg are limited to those in the public domain, which means Western classics that are no longer (or never were) protected by copyright make up the majority of the e-books it provides. For the most part, anyone can search and read the books for free of charge without having to use a special application or program.<sup>12</sup> In addition, e-books are offered not only in English but also in French, Portuguese, and German. On occasions Project Gutenberg makes requests to copyright holders to share even more books with the world.

A notable characteristic of Project Gutenberg is that anyone and everyone can have access to most classical literary works without constraints. Unlike most e-books, whose formats differ depending on the publisher or device, all works available on the Project Gutenberg site are provided in the HTML format, as well as in e-pub and other formats for e-reader access. In addition, Project Gutenberg does not claim copyrights for these e-books and actively supports their use in various ways. Another important trait of this project is that their e-books are polished through proofreading and revisions and formatted according to their own standards.

The development of e-books for Project Gutenberg involves the use of an innovative technology called Distributed Proofreaders. This is an early example of crowdsourcing, which allows volunteers to work together and

---

ject\_Gutenberg\_by\_Michael\_Hart

11 This information is from Greg Newby's presentation entitled "Forty-Five Years of Digitizing eBooks: Project Gutenberg's Practices" at the 15th International Workshop for Translation and Publication of Korean Literature (2016).

12 This is also a result of one of the Project Gutenberg's principle: No Fees, and No Custom Apps Required.

**Project Gutenberg**

search for books

- Browse Catalog
- Bookshelves
- Main Page
- Categories
- Contact Info

Project Gutenberg appreciates your donation!

[Donate](#)

- Why donate?

In other languages

- Português
- Deutsch
- Français

[ibiblio](#)

## Free eBooks - Project Gutenberg

[Book search](#) · [Book categories](#) · [Browse catalog](#) · [Mobile site](#) · [Report errors](#) · [Terms of use](#)

### Some of the Latest eBooks

**Welcome**

**Project Gutenberg** offers over 58,000 free eBooks. Choose among free epub and Kindle eBooks, download them or read them online. You will find the world's great literature here, with focus on older works for which U.S. copyright has expired. Thousands of volunteers digitized and diligently proofread the eBooks, for enjoyment and education.

**No fee or registration!** Everything from Project Gutenberg is gratis, libre, and completely without cost to readers. If you find Project Gutenberg useful, please consider a small [donation](#), to help Project Gutenberg digitize more books, maintain our online presence, and improve Project Gutenberg programs and offerings. Other ways to help include [digitizing](#), [proofreading and formatting](#), [recording audio books](#), or [reporting errors](#).

[Project Gutenberg Mobile Site](#)

Fig. 4. Main Page of Project Gutenberg Website

participate in the development of e-books by taking care of copyright issues, proofreading, formatting, reviewing, and making a final check before the e-books are uploaded to the site. The organization of these processes has made it easier for print materials to be digitized and developed into e-books.

Project Gutenberg has spearheaded the sharing of e-books in an unprecedented manner and currently provides access to standardized e-books without copyrights that anyone can use, owing to the efforts to volunteers. It has made groundbreaking achievements in allowing ease of access to the classics and has become the first “virtual library.”

There are also projects that make use of classics as historical sources instead of simply providing access to them as e-books. One example is the Chinese Text Project (C-text).<sup>13</sup> A collection of various Chinese historical sources, it was created by Donald Sturgeon, who majored in East Asian Language and Civilizations at Harvard University.

<sup>13</sup> <http://ctext.org>

The screenshot shows the homepage of the Chinese Text Project. At the top, there are social media icons for Facebook, Twitter, and Weibo. The main heading is 'Chinese Text Project' with the Chinese characters '百諸家子' below it. A 'Welcome' section follows, explaining the project's mission as an online open-access digital library. Below this, there are links to 'read more about the project', 'pre-Qin and Han', 'post-Han', and 'Wiki' tables of contents, and a link to 'consult the instructions, FAQ, or list of tools'. A search bar is provided for finding specific texts. The 'Latest additions' section is partially visible at the bottom. The left sidebar contains a detailed navigation menu with categories such as 'About the site', 'Qin and Han', 'Post-Han', and various sub-topics like 'Confucianism', 'Mohism', 'Daoism', etc.

Fig. 5. Main Page of the C-Text

First founded as part of an attempt to examine how studies in different fields are connected to each other, C-text provides access to textual data, which currently amounts to a total of 26,072,942 Chinese characters. C-text has been transcribing scanned images of source texts, and 7.9 percent of the materials have been transcribed. One thing to note is that this project is not undertaken by Donald Sturgeon alone. Over 30,000 users visit the site daily, and they not only browse the site but also contribute to the site by digitizing texts in print and linking them to other relevant texts. In other words, there is no barrier between users and providers.

Another major characteristic of C-text is that it is easily accessible. Administrators communicate with users via message boards, and the Discussion and Wiki sections on the site are particularly useful. The Discussion message boards provide a platform for users to discuss textual issues, such as Chinese characters that are difficult to decipher or errors in different editions of historical sources. In some circumstances, it can lead to a consuming debate between non-experts, but in most cases these message boards allow people to use their collective knowledge to find the right answers. For instance, on November 28, 2018, a member with the username “oscarsun 72” correctly pointed out the missing characters in the OCR output for *Shiji Sanjiazhu* (Records of the Grand History with commentaries of the three experts), and C-text incorporated the user’s

opinion and rectified the issue.

The Wiki section is even more interesting. The number of historical sources for Chinese history is much too vast for individuals to grapple with. In addition, people have historical sources in various formats—some have photographic images of sources, while others have them in the text form. In the Wiki section, C-text allows users to share their sources. A user who has files of historical sources that are yet unavailable on C-text can upload them onto the Wiki section, providing access to those who do not have those sources. Already, over 10,000 volumes of texts have been submitted by users.

The success of C-text was not simply built on individual efforts. In 2015, the Wenyuange Edition of *Siku Quanshu* (Complete Library of the Four Treasuries, the largest collection of texts in Chinese history) was converted into text using OCR (optical character recognition) on C-text, and, in 2016, the Harvard-Yenching Library provided over 50,000 pages of resources. Everyone who used C-text participated in reviewing these sources, redressing the errors resulting from the lack of accuracy in the OCR process. As a result, C-text has now grown into an online *Siku Quanshu*.

C-text cannot replace all historical sources. Ultimately, it is up to the researchers to directly assess historical materials, while online sources should be used as secondary sources. However, it is true that these online sources have made it much easier for scholars to conduct research and have even revealed things that were difficult to identify in the past. In his research, Sturgeon, who created C-text, analyzed user-created sources on C-text and identified the methods and links between text reuse in early Chinese literature.

Various attempts have also been made for using C-text in education. Harvard and Duke universities in the United States, as well as Japan's Kyoto University, China's Tsinghua University, and Taiwan's National Taiwan University are already working to incorporate C-text into their educational programs. Since C-text allows scholars to compare and analyze a vast amount of Chinese historical sources, universities seem to be



NIKH's database provide the same original text of *Koryŏsa*, but the two are completely separate platforms, operated by different entities. As a result, errors in the descriptions or words that are difficult to decipher cannot be corrected on both sites at once, as the operators of the two sites have to be alerted separately. In addition, while the NIKH's database has become much easier and more convenient to use due to the results of recent projects, it is becoming difficult to use Naver's service due to the integration with other digital encyclopedias.

Research on the Koryŏ dynasty has become easier and more technologically advanced. The biggest advantage Koryŏ dynasty researchers have is the revised chronology and tag words, which have made it possible for scholars to view all historical sources related to a certain date or a certain historical figure in a glance. However, this is only possible on the exclusive Research Institute Goryeosa website, and users have to use other services to look at Korean translation or photographic images of the original texts.

There is another problem for these digital archives—there are a lot of resources that are not available on these platforms. For instance, there are different editions of *Koryŏsa*, and there are discrepancies between the different editions. While the comparison of different editions of texts is a crucial factor in research, websites that provide access to historical sources tend to provide only one version. Yet another issue is in the discrepancy in the level of services provided by platforms. While it would be difficult for all providers of sources from Koryŏ dynasty to provide access to all Koryŏ dynasty sources online, it is problematic that some materials that are easily accessible in print format, such as epitaphs and *Xuanhe fengshi Gaoli tujing*, are only available on a select few platforms.<sup>14</sup>

---

14 For the digitization of historical sources from the Koryŏ dynasty, see Ch'oe Yŏnju, "Koryŏ sidae mijŏngbohwa charyo-ŭi hyŏnhwang-gwa chŏnsanhwa pangan: myojimyŏng, ilban kŏmsŏngmun-ŭl chungshimŭro" (The status and digitization method of non-digitized historical sources from the Koryŏ dynasty: focusing on

Aside from the examples mentioned in this paper, there have been other attempts to digitize historical sources from the Koryŏ dynasty. The Tripitaka, miscellanies, and historical manuscripts from the Koryŏ dynasty have mostly been digitized,<sup>15</sup> and databases of sources on a certain topic, such as the island database, have been created and made accessible as well.<sup>16</sup> These services, however, are also limited in that they only provide access to sources that are relevant to a particular topic. This is because while each site was created for all Koryŏ dynasty researchers, it is not the researchers who come together to manage the sources themselves. Instead, only a few are responsible for taking care of the sites.

On the contrary, online services in foreign countries that provide access to historical sources mostly took a different approach. As seen in Project Gutenberg, e-books of various manuscripts are being freely shared and distributed. An important point of note is that most of these e-books are

---

epitaphs and general epigraphs), *Han'guk chungsesa yŏn'gu* (The Journal of Korean Medieval History) 30 (April 2011). This paper, however, presents the state of affairs from a decade ago. Since then, considerable progress has been made in the digitization of epitaphs and epigraphs.

- 15 Ch'oe Yŏnggho, "Koryŏ sidae taejanggyŏng, munjip, komunsŏ charyo-ŭi chŏngbohwa hyŏnhwang-gwa chŏnsanhwa pangan" (A study on the digitization of Tripitaka miscellanies, and historical manuscripts from the Koryŏ dynasty), *Han'guk chungsesa yŏn'gu* (The Journal of Korean Medieval History) 30 (April 2011).
- 16 Hong Yŏngŭi, "Han'guk yŏksasang ko, chung-se tosŏ (sŏm) kwallyŏn kich'ojarjo DB kuch'uk-gwa hwaryong: tosŏ (sŏm) yŏksa, munhwa, chirihak-ŭl t'onghan haeyanggwŏnyŏk yŏn'gu hwalsŏnghwa" (Construction and utilization of a database of basic data on islands in ancient-medieval periods in Korean history: vitalization of research on the territorial waters of Korea through the study of island history, culture, and geography), *Han'gukhak nonch'ong* 46 (September 2016).
- Hong Yŏngŭi, 2018, "Han'guk sŏm-gwa yŏnan-ŭi munhŏnjaryo teit'ŏ(DB)-ŭi kuch'uk-gwa hwaryong chŏnmang (The construction and expected use of a database of historical sources on Korean islands and coastal areas), *Han'gukhak nonch'ong* 50 (August 2018).

provided in formats that anyone can use. C-text is operated in a similar way, as a host of users provide data that can be shared, increasing the size and scope of the database. In addition, both Project Gutenberg and C-text are open-source collectives, where all users come together to manage the sources and contribute to the site. The impact of such collaborative work is particularly evident in C-text, as vast amounts of sources are being organized within a short period of time. This would have been difficult to be achieved by a single individual or a single organization.

These examples provide important implications for the direction of research on the Koryŏ dynasty, particularly for the organization of historical sources for research. As mentioned earlier, a big shortcoming of the digitization of historical sources from the Koryŏ dynasty is that the different platforms and services that provide access to those historical resources are not connected. Moreover, the sources provided have been created and refined through the efforts of certain groups of people. In the future, however, this kind of exclusivity should be reconsidered and removed. The first task for remedying this issue would be to create a single, unified platform for all historical sources. Major historical sources from the Koryŏ dynasty are *Koryŏsa* and *Koryŏsa chŏryo*. Additionally, *Xuanhe fengshi Gaoli tujing* and epitaphs are used as primary sources for research on the Koryŏ dynasty as well. Furthermore, *Cholgo ch'ŏnbaek* (a collection of literary works by Ch'oe Hae), *P'ahanjip*, and other collections of literary works from Koryŏ and later periods also mention the events that have taken place during the Koryŏ dynasty. Providing all of these sources via a single platform will be difficult as copyright and other matters have to be resolved. But none of these resources can be cast aside in the study of the Koryŏ dynasty. As mentioned earlier in this paper, the absolute number of sources is very small for the Koryŏ dynasty compared to other periods, which makes every single resource ever so important. Therefore it would be meaningful to create a single platform that gives access to all of these sources to be compared and analyzed together.

Various historical sources are used in research today. But most collections of literary works have not even been translated and certainly not

digitized. Converting these sources into texts and furthermore translating them will be valuable efforts, which would benefit from the participation of more people. While there may be qualitative differences in translations depending on the translators' understanding of classical Chinese and nuances, conversions of images into texts would simply require a similar level of efforts put into C-text. A single organization or collective's efforts to examine and review numerous Chinese characters would not be enough to produce successful results.

The merit of C-text is not that it is a huge repository of Chinese historical sources or that it is easy to use. Rather, its asset is the results of collective knowledge, gained by numerous users proofreading the sources and exchanging opinions. Images of documents can be converted into texts through technology. However, improving the quality of sources by going over the converted texts and correcting the errors is impossible without human effort. C-text made this possible by providing a platform and eliciting people's voluntary participation, which has been producing successful results.

Can the organization of historical sources from the Koryŏ dynasty also head toward this direction? Unlike *Koryŏsa* and *Koryŏsa chŏryŏ*, which have been studied and verified by researchers numerous times in the past, many of the collections of literary works from Koryŏ have not yet undergone such verification processes. In addition, extracting contents relevant to the Koryŏ dynasty from extensive historical sources from the Chosŏn dynasty and organizing them will be a daunting task in itself. If we can open up the process of such tasks to create texts with better quality, the results will be of invaluable significance to the study of the Koryŏ dynasty.

In consideration of these circumstances, the first objective that a future platform for historical sources from the Koryŏ dynasty should pursue is openness. A vast number of sources are already available online. If we can gather these sources together on a single platform and secure numerous contributors, we would be able to say that openness has been achieved for Koryŏ dynasty sources. Contributors will be more than just

users, as they will make improvements to current technological limitations in OCR output. Once openness is secured, not only scholars but also the general public will be able to easily access various historical sources and compare them. This will not only bring openness to research on the Koryŏ history but also lead to the Koryŏ history itself.

It is important to clarify that too much openness can disrupt and hinder historical research. Therefore the second objective that a future platform for historical sources from the Koryŏ dynasty should pursue is expertise. If a historical resource platform is open to all, anyone can have access to historical sources. The platform, however, should not stop at simply providing access to texts. Instead, it should take a step further to work on establishing a revised chronology using the given sources and engage in metadata tagging, as these are essential elements for a future-oriented organizational method for historical sources.

The revision of the chronology and tagging needs to be based on a general organization of the sources by research experts. The understanding of the Koryŏ dynasty and the skills to analyze historical sources are crucial to these tasks. This is where “expertise” comes in. It is an area that requires research experts who have acquired the basic skills to conduct historical research through education.

The importance of expertise in studying the Koryŏ dynasty is particularly high when considering the research environment. Historical sources from the Koryŏ dynasty are small in number compared to other periods, and therefore an extensive understanding of available sources requires high precision and accuracy. In addition, the number of researchers who will be reproducing and reusing the sources is also relatively small, which means that participation of research experts is necessary to limit the damage that openness of the service may inflict on the quality of sources.

In this paper, I have examined the direction that a future digital archive platform should pursue in order to provide historical sources from the Koryŏ dynasty and arrived at the conclusion that “openness” and “expertise” are the answer. Providing easy access to historical sources and allowing participation from all users with the guidance of research experts

will be a way to bring about both an increase in the public's interest in the study of the Koryŏ dynasty and ensure ease and convenience in research. Among the various nations and dynasties that have existed on the Korean Peninsula, Koryŏ was the most open and, at the same time, most complex one. It is now time to incorporate both of these attributes—openness and complexity—to the future of research on the Koryŏ dynasty.

## References

1. Ch'oe, Yŏngho, "Koryŏ sidae taejanggyŏng, munjip, komunsŏ charyo-ŭi chŏngbohwa hyŏnhwang-gwa chŏnsanhwa pangan" (A study on the digitization of Tripitaka miscellanies, and historical manuscripts from the Koryŏ dynasty), *Han'guk chungsesa yŏn'gu* (The Journal of Korean Medieval History) 30 (April 2014).
2. Newby, Gregg, "Forty-Five Years of Digitizing eBooks: Project Gutenberg's Practices" at the 15th International Workshop for Translation and Publication of Korean Literature 2016
3. Hong, Yŏngŭi, "Han'guk yŏksasang ko, chung-se tosŏ (sŏm) kwallyŏn kich'ojaryo DB kuch'uk-gwa hwaryong: tosŏ (sŏm) yŏksa, munhwa, chirihak-ŭl t'onghan haeyanggwŏnyŏk yŏn'gu hwalsŏnghwa" (Construction and utilization of a database of basic data on islands in ancient-medieval periods in Korean history: vitalization of research on the territorial waters of Korea through the study of island history, culture, and geography), *Han'gukhak nonch'ong* 46 (September 2016).
4. Hong, Yŏngŭi, "Han'guk sŏm-gwa yŏnan-ŭi munhŏnjaryo teit'ŏ(DB)-ŭi kuch'uk-gwa hwaryong chŏnmang (The construction and expected use of a database of historical sources on Korean islands and coastal areas), *Han'gukhak nonch'ong* 50 (August 2018).
5. Pak, Chinhun, "Koryŏ sidae munhŏn charyo chŏngbohwa saŏp hyŏnhwang mit iyong shilt'ae-wa hyogwajŏgin hwaryong bangan (Projects for the digitalization of historical documents from the

- Koryŏ dynasty, their use and effective utilization method),” *Han’guk chungsesa yŏn’gu* (The Journal of Korean Medieval History) 30 (April 2014)
6. Shin, Ŭnje, “*Kugyŏk Koryŏsa*-ŭi tochŏn, kŭrigo han’gye (Challenges and limits of *Kugyŏk Koryŏsa*),” *Sŏkdang nonch’ong* (Journal of Seokdang Academy) 54 (November 2012)
  7. Song, Yongdŏk, 2018, “Kuksap’yŏnch’an wiwŏnhoe Koryŏ sidae saryo chŏngpohwa hyŏnhwang-gwa kwaje” (The status and challenges of the National Institute of Korean History’s digitalization of historical sources from the Koryŏ dynasty), *Koryŏ kŏn’guk 1,100 chun-yŏn kinyŏm che 53 hoe Kuksap’yŏnchan wiwŏnhoe han’guksa haksurhoeŭi kyŏnggye-rŭl nŏmŏ saeroun kilo* (The 53rd National Institute of Korean History academic conference in celebration of 1,100 years since the founding of the Koryŏ dynasty: a new path beyond the boundary), (Seoul: Kuksap’yŏnchan Wiwŏnhoe, 2018)
  8. Yi, Namhŭi, “*Koryŏsa* tijit’ŏlhwa-ŭi panghyang-gwa kwaje” (The direction and challenges of the digitization of *Koryŏsa*), *Ch’ŏnggyesahak* 18 (August 2003)

〈Abstract〉

## **A New Path for the Study of the Koryŏ Dynasty: Exploring the Future of Online Historical Source Archives**

*Soochan Park*

The first objective that a future platform for historical sources from the Koryŏ dynasty should pursue is openness. A vast number of sources are already available online. If we can gather these sources together on a single platform and secure numerous contributors, we would be able to say that openness has been achieved for Koryŏ dynasty sources. Contributors will be more than just users, as they will make improvements to current technological limitations in OCR output. Once openness is secured, not only scholars but also the general public will be able to easily access various historical sources and compare them. This will not only bring openness to research on the Koryŏ history but also lead to the Koryŏ history itself.

It is important to clarify that too much openness can disrupt and hinder historical research. Therefore the second objective that a future platform for historical sources from the Koryŏ dynasty should pursue is expertise. If a historical resource platform is open to all, anyone can have access to historical sources. The platform, however, should not stop at simply providing access to texts. Instead, it should take a step further to work on establishing a revised chronology using the given sources and engage in metadata tagging, as these are essential elements for a future-oriented organizational method for historical sources.

**Keywords:** Koryŏ, Koryŏsa, Database, Online, Historical Source Archives, openness, expertise

〈국문초록〉

## 고려시대 연구의 새로운 길 - 사료 서비스의 미래를 중심으로

박수찬 (고려대학교 박사과정)

미래의 고려시대 사료 서비스가 추구해야할 첫 번째 길은 바로 ‘개방성’이다. 이미 사료들은 온라인 상에서 수없이 제공되고 있다. 이들을 하나로 묶어 정리하고, 이들에 대한 수많은 공여자를 확보할 수 있다면 고려시대 사료의 개방성이 확보된 것으로 파악할 수 있을 것이다. 또한 그 공여자들은 단순히 사료를 활용할 뿐만 아니라 OCR 등의 작업에서 이루어진 기계적 한계를 잡아내는 역할도 할 수 있다. 이처럼 개방성이 확보된다면 연구자들 뿐만 아니라 고려시대에 관심있는 수많은 일반 대중들 역시 다양한 사료를 더 쉽게 접하고 비교할 수 있게 될 것이다. 이는 비단 연구의 영역 뿐만 아니라 고려시대 역사 그 자체가 개방성을 확보할 수 있는 길이 되는 작업이다.

그러나 지나친 개방성은 오히려 역사 연구의 방해 요소로 작용할 수 있다는 점 역시 명확히 해야 한다. 이를 위해서 고려시대 사료 서비스가 추구해야할 두 번째 길은 ‘전문성’이다. 텍스트화한 자료에 대한 수정 편년 작업, 태깅 작업 역시 연구의 용이성을 확보하기 위해서는 꼭 갖추어야할 요소라고 할 수 있겠다. 이러한 수정편년이나 태깅 작업의 경우 전문화된 연구원 중심으로 기본적인 정리가 우선적으로 이루어질 필요가 있다. 이는 고려시대 전체에 대한 이해와 사료를 분석할 수 있는 능력을 갖추지 않았다면 정확한 선택에 어려움이 있기 때문이다. 그러므로 필요한 것이 바로 ‘전문성’이다. 전문적인 교육을 통해 역사 연구를 할 수 있는 기본 소양을 갖춘 전문화된 연구원이 필요한 영역인 것이다.

**주제어:** 고려, 고려사, 데이터베이스, 온라인, 사료 서비스, 개방성, 전문성

