Case Study on Sharing and Cooperation on East Asian Digital Resources : Focusing on the East Asia Digital Library(EADL) pilot service

Haeryoun Kim

Digital Library Planning Division, National Library of Korea, Seoul, Republic of Korea
hrkim15@korea.kr

Abstract

In December 2020, the National Library of Korea(NLK) and the National Diet Library(NDL) of Japan launched a pilot service of the East Asia Digital Library(EADL). About 8,000 digitized materials can be used freely by anyone online via EADL. These materials have been digitized over the years by NLK and NDL for users in their countries. At EADL, these materials can be accessed using integrated search and curated collection browsing. This is likely to be a very useful service for various users such as researchers who need to use the materials of Korea and Japan. This study analyzed the contents and services provided by the EADL pilot service and discussed ways to improve them.

Keywords: *Digital Library, East Asia, National Library of Korea, National Diet Library*

1. Introduction

Over the past 30 years, many libraries have digitized their collections. As a result, libraries have a fairly large amount of digital resources. Due to this, effective digital library services have recently become available. By simply accessing the digital library, digital resources can be used by anyone, anytime, anywhere in the world. This is especially the case for historical materials that have no copyright restrictions. For these materials, it is possible to provide easy access and various methods of use so that anyone can use the digitized materials. It is very meaningful for various institutions and countries to work together in order to provide these services. World Digital Library and Europeana are representative examples of international cooperation in digital libraries. Libraries, museums and archives from various countries are participating in the World Digital Library and the Europeana. Thanks to these large projects, the digital cultural heritage and digital knowledge resources available on the Internet have increased. In addition, researchers who study history and culture and art can easily access a wealth of research materials.

The National Library of Korea(NLK) has digitized over 8 million materials and provides them online. It would be very meaningful to provide these digitized materials to users not only in Korea but also around the world. The National Diet Library(NDL) of Japan also provides more than 2.75 million digitized materials. Among them, there are many materials whose copyright protection period has expired, and these materials can be used for free by anyone on the Internet. If these digital resources of Korea and Japan are provided to anyone around the world to use online, it will be a useful service for researchers doing research on Korea, Japan and East Asia, as well as students and general users interested in East Asia. In March 2020, NLK and NDL signed an MOU to share digital resources and provide new access. Since then, NLK and NDL have been cooperating to build an East Asia Digital Library(EADL). As a result, in December 2020, the EADL pilot service was launched. The EADL is an online service that provides integrated search

and access to the digital collections held by NLK and NDL. Over 8,000 digital materials and metadata are available via EADL. EADL is a pilot service stage, so there are many points to be improved in the future. More institutions need to participate in EADL and more digital resources must be provided via EADL. In doing so, EADL should be a service that is actually helpful to users. This study analyzes the contents and services of the EADL pilot service and proposes improvement plans.

2. EADL contents

2.1. Scope and features

Currently, EADL allows users to access 119 digitized books and 3,926 metadata from NLK's digitized materials. Users can also access NDL's 80 digitized books and 3,988 metadata(Table 1).

Table 1. Current status of data available on EADL

Institution	Digitized books	Metadata	Total
National Library of Korea	119	3,926	4,045
National Diet Library of Japan	80	3,988	4,068
Total	**199**	**7,914**	**8,113**

These are ancient and modern materials published between 700 and 1945, and the era distribution of the materials is quite wide. Table 2 shows the distribution of materials by publication year. Since EADL's materials are mainly old books, there are many materials whose publication year is unknown. Except for the materials of unknown publication year, the materials provided by the NLK were mainly published after 1800. Many of the materials provided by the NDL were published between 1600 and 1800.

Table 2. Current status of materials by publication year

Institution	Unknown year	~1000	~1200	~1400	~1600	~1800	~2000	Total
NLK	1,238	0	3	25	349	332	2,098	4,045
NDL	1,874	2	9	35	244	1,894	10	4,068
Total	**3,112**	**2**	**12**	**60**	**593**	**2,226**	**2,108**	**8,113**

Table 3. Current status of materials by chronicle

Korean chronicles		NLK materials	NDL materials	Japanese chronicles	
Unified Silla	(676-935)	0	2	Nara era	(710-794)
			6	Heian period	(794-1185)
Goryeo Dynasty	(918-1392)	22	15	Kamakura shogunate	(1185-1333)
			116	Muromachi shogunate	(1336-1573)
Joseon Dynasty	(1392-1897)	1,007	161	Azuchi Momoyama	(1568-1603)
			1,892	Edo shogunate	(1603-1868)
Korean Empire	(1897-1910)	143	2	Japanese Empire	(1868-1945)
Japanese occupation	(1910-1945)	768			
Republic of Korea	(1945-)	867	0	Japan	(1945-)
Unknown		1,238	1,874	Unknown	
Total		**4,045**	**4,068**		

On the other hand, Table 3 shows the status of materials by each country's chronology. Except for the materials of the unknown period, there are many materials from the Joseon Dynasty(1392-1897) among Korean materials. And there are many Japanese materials from the Edo period(1603-1868).

Table 4 shows these materials by language. There are many materials written in Chinese characters among the NLK and NDL materials. This shows that both Korea and Japan belong to the Chinese character culture.

Table 4. Current status of materials by language

Insti.	Korean	Chinese	Japanese	English	Dutch	Germany	Russian	Manchu	Combodian	Total
NLK	95	3,939	27	3	0	1	1	1	0	4,067
NDL	54	1,197	2,814	0	2	0	0	0	1	4,068
계	**149**	**5,136**	**2,841**	**3**	**2**	**1**	**1**	**1**	**1**	**8,135**

Still, there are not many materials available on EADL. However, if you look at the composition of the current materials, you can see the service scope and features of EADL. EADL can be said to be a service to share and provide access to historical materials of East Asia, which belong to the Chinese character culture region, such as Korea, Japan and China, so that they can be used anywhere in the world. The scope of materials provided by EADL is not limited to those held by institutions in the East Asia region. The purpose is to share materials published in East Asia such as Korea, Japan, China, etc. or written by authors in East Asia or related to East Asia. Therefore, any institution that has East Asian materials can be a participating institution in EADL, regardless of whether it is in the United States, Europe or anywhere in the world. As is the case with materials from other cultures, East Asian old materials are held by various institutions around the world. If such dispersed materials can be searched and accessed online, it would be very useful to related researchers.

2.2. Data construction and management

For EADL service, NLK and NDL provided metadata and original image files of digitized materials. In EADL, provided metadata is saved in Linked Data format and managed according to RDF-based ontology rules. EADL's data model is shown in Figure 1. CJKObject means data describing the original printed materials of digitized materials, and DigitizedObject means data describing the digitized materials. Also, ImageObject is data describing digital image files of digitized materials. Collection is data describing a digital collection composed of digitized materials. In the EADL data model, DigitizedObject and ImageObject are connected and managed centering on CJKObject, and CJKDLCollection is formed by organizing them.
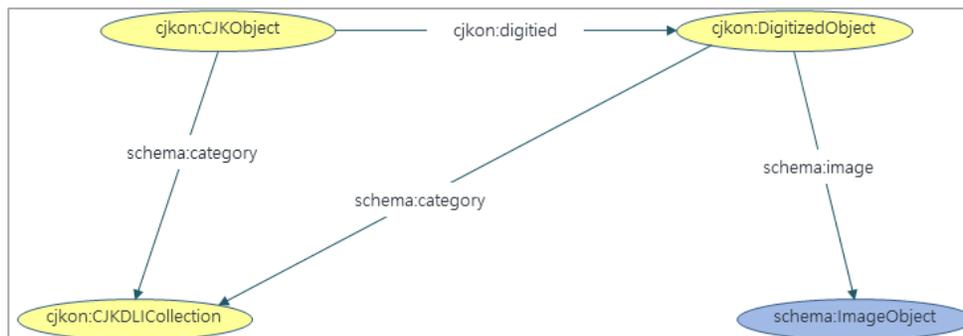

Figure 1. EADL ontology model

The metadata items for each of these types are shown in Table 5. In order to secure global interoperability, a metadata item based on DublinCore was constituted. The final saved data format is shown in Figure 2.

Table 5. EADL metadata item

Type	Term (mandatory)	Term (optional)	Note
Collection	Country / URI(Depth1)~ URI (Depth3) / Label / Description		Description of the Collection
Object	ID / Category / Label / Country that provides data / DataProvider	Author / Publisher / Subject / date / Issued Date / Description / Abstract / Language / Alternative title / Audience / Size / Publication Place / Local Holding and so on	Description of the original materials
DigitizedObject	ID / Category / Format / Publisher who digitized content / Holding Agent	Volume Number / Source / Thumbnail / Next Volume / URL viewing contents / Rights / contributor / Edition / Publisher / DOI / Date digitized / Medium / Number of image file	Description of the digitized materials
ImageObject	ID / Volume Number / Locator / Format / Bit-depth / File Path	Color Space / Compression Ratio / Resolution / Scan Ratio / Original Source	Description of the digitized image files

Figure 2. LOD publication page for EADL materials

In the case of objects with digitized images, major metadata such as Title, Creator, Publisher and Abstract are provided in three languages: English, Korean and Japanese. Digitized images of materials are provided to be viewed, printed and downloaded via the EADL viewer or viewers of each data provider(Figure 3).
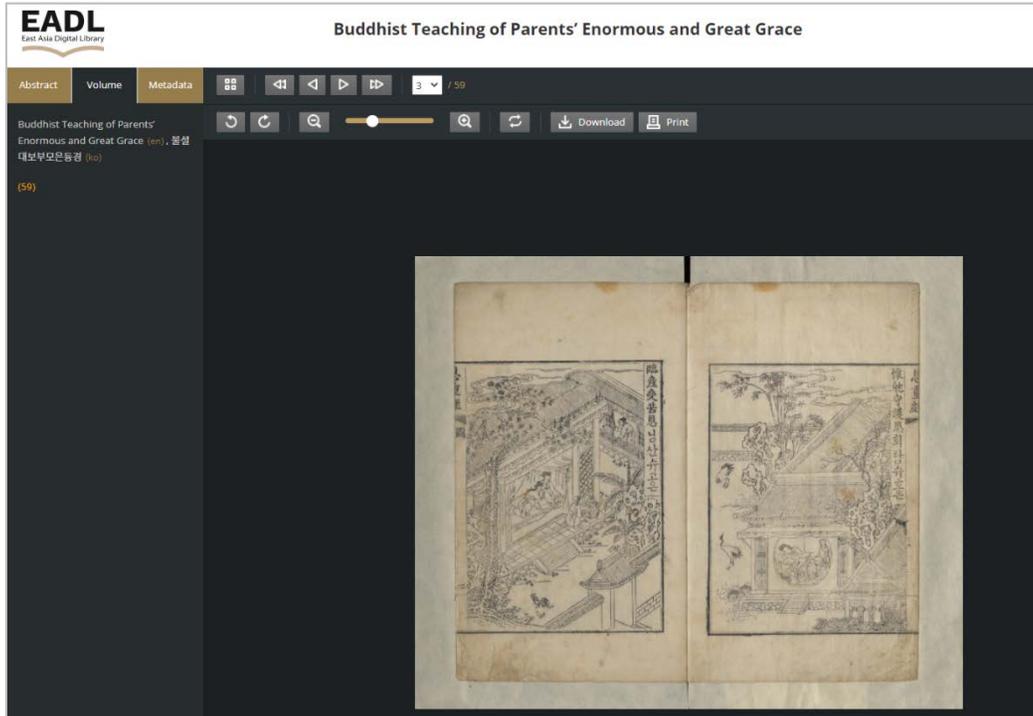

Figure 3. EADL viewer

3. EADL services

3.1. Target users

The target users of EADL are researchers who study East Asian Studies, Korean Studies, Japanese Studies and Chinese Studies. The greatest meaning of EADL is that it can provide digitized materials to researchers who explore East Asian history, culture, thought, literature and art. Also, for students and general users interested in East Asia, EADL can be a gateway to find useful materials. EADL will provide East Asian copyright expiration materials that can be freely captured or downloaded over the Internet. Then, users will be able to use EADL's materials for their own content or share it on social media. This will help users know about East Asia and increase their interest in East Asia. It will also encourage active use of East Asian materials.

3.2. EADL website

The EADL website supports the use of materials on the web. Figure 4 is the main page of the EADL website. The menu of the website is shown in Table 6. EADL is basically provided in English. However, some textual content such as About, FAQ, Terms and Policies are also available in Korean, Japanese and Chinese.
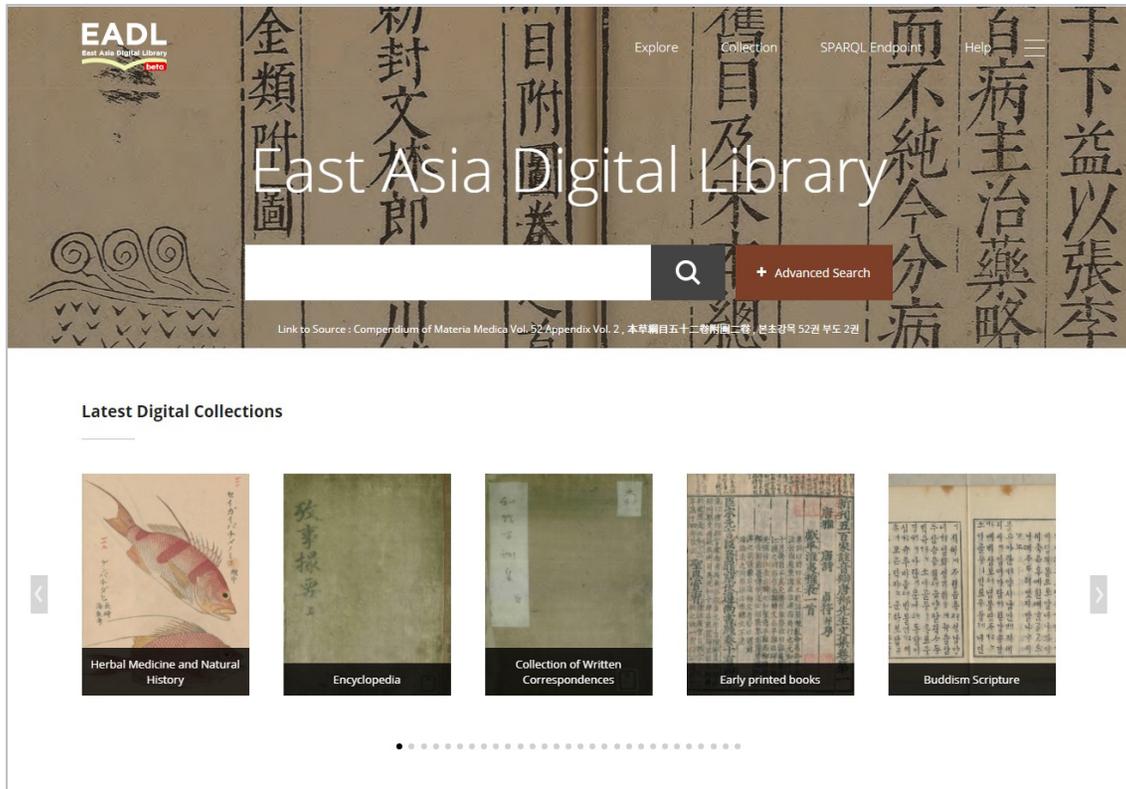
Figure 4. Main page of EADL website

Table 6. Menu of EADL website

Menu	Sub-menu	Function
Explore		It shows the full list of EADL materials
Collection	By Topic	It shows digital collection materials by topic
	By Institution	It shows digital collection materials by institution
	By Time	It shows digital collection materials by time
	By Chronicle	It shows digital collection materials by chronicle
SPARQL Endpoint		It provides SPARQL query function
Help	FAQ	It provides frequently asked questions and answers
	System Documents	It provides service information, ontology terms and ontology specifications etc.
Bottom information	Information	Introduction of EADL and participating institutions, contact information, Terms and Policies
	Contact Info	Contact information of EADL
	Quick Link	Links to NLK, NDL and SPARQL endpoint

The search box is fixed at the top of the website so that you can use the search function whatever page you are on. You can search by keywords of title, author or publisher etc., but not by keywords in Abstract. In addition, facets are provided so that the search results can be limited by holding institution, region, material type, language and publication year.

Collection is a service that enables users to browse materials by subject, institution, time and chronology. In 'By topic', materials are collected and provided by subject or type of material such

as Korean arts, Korean Buddhism, Japanese old printed books, herbal medicine and natural history(Table 7). In 'By institution', you can see the collections of NLK and NDL that are currently provided. In 'By time' you can browse materials by publication year and in 'By chronicle' you can browse the materials by chronology(Figure 5).

Table 7. EADL collection by topic

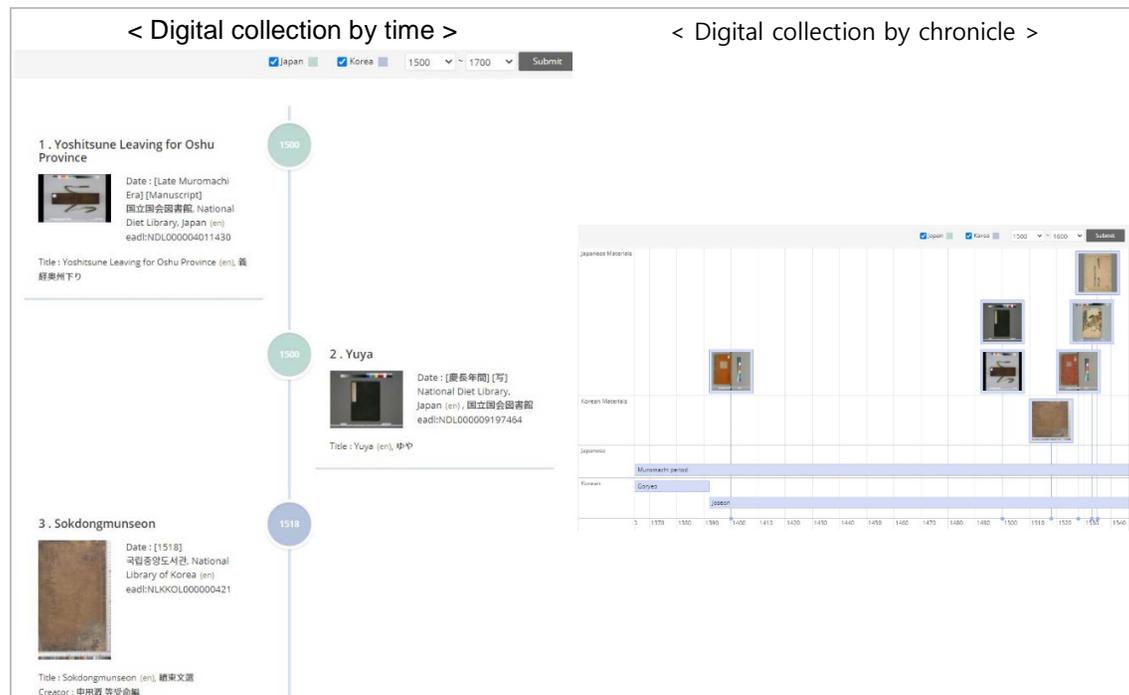| Institution	Topic	Digitized materials
NLK	Documents of the Joseon Royal Family(25) / Korean arts(12) / Korean Buddhism(8) / Joseon Tongsinsa (21) / Women in Joseon era(18) / Dictionaries from Joseon Dynasty(5) / Scenic spots with beautiful landscapes(8) / First paperback novels in Korean, Ttakjibon(22)	119
NDL	Herbal medicine and natural history(10) / Old movable-type printed books(11) / Old printed books(9) / Others(50)	80

Figure 5. EADL browsing by time and chronicle

3.3. Data services

Linked Data is structured data which is interlinked with other data. It builds on standard web technologies such as HTTP, RDF and URIs. However, rather than using data to serve web pages only for human readers, Linked Data extends data to share information in a way that can be read automatically by computers. The vision of Linked Data is to make Internet a global database(by Wikipedia).
SPARQL is the standard query language of W3C to extract desired data from such Linked Data. As SQL is a standard query language in relational databases, SPARQL is used in Linked Data. Being able to query SPARQL means that it provides developers or users with a higher degree of freedom to access data conveniently. SPARQL Endpoint is an access point at which you can make SPARQL queries on the web(by NLK LOD).
EADL data is saved and managed in Linked Data format. If SPARQL Endpoint is provided, users can directly access data published as Linked Data through query. Therefore, EADL provides

SPARQL Endpoint from a data service point of view. Figure 6 shows an example of using SPARQL Endpoint. Users can download data not only in RDF/XML format, but also in JSON, Turtle and other formats.
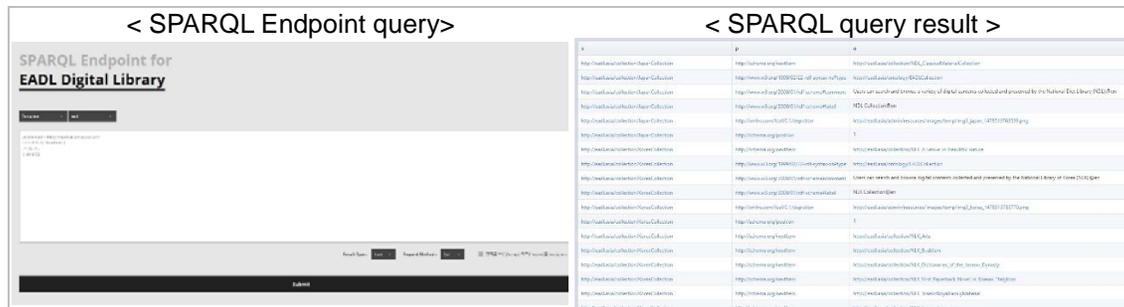

Figure 6. Example of using SPARQL Endpoint

4. Analysis and suggestion

Currently, EADL is open as a beta version and is in pilot service. Official service will begin in the future by improving errors or problems found through the pilot service and reflecting opinions of users. In this study, contents and services of EADL pilot service were analyzed. The results are as follows.

First, EADL could be a very useful service for researchers in related fields as EADL is a service that enables integrated search and use of digitized materials of national libraries in Korea and Japan. However, not enough materials have been collected yet to satisfy users. In fact, in order to become a useful service for researchers there needs to be much more materials available on EADL. Also, EADL currently provides only materials from Korea and Japan, so the range of materials provided is narrow compared to the name 'East Asia'. In addition to materials of Korea and Japan, it is necessary to collect and provide various materials from Chinese character culture regions including China. Then, the service that EADL aims for will become possible.

Secondly, EADL's target user setting is highly efficient. Since there are similar digital library services provided by various institutions, in order for EADL to differentiate its service, it is necessary to clarify the scope and features of materials provided and clarify target users. To do this, data collection and service planning should be done based on target user demand investigation. That way users can request a variety of services by research field, language and purpose of use. During the pilot service period, the user demand analysis and detailed service planning needs to be carried out.

Thirdly, EADL allows access to data through basic search function and search result restriction function through facets. The facets fit the features of materials provided on EADL. However, the search function needs to be continuously improved over the long term. It is necessary to make it convenient and easy to access data through searching by securing and utilizing controlled vocabulary. In order to improve the search function, EADL should consider the characteristics of Chinese characters and East Asian materials such as traditional and simplified Chinese characters, Japanese Chinese characters, Korean notation, Japanese furigana and Chinese pinyin. In addition, in EADL, metadata is not used as a path value to other materials, but only as the data itself. If metadata is used as a path value, users can move to view materials of the same title, author, era, subject etc.

Fourth, EADL provides data services such as SPARQL Endpoint for users to use data usefully. In the future, it will be possible to provide more various types of data besides Linked Data. Data service will become an important service not only to ordinary users but also East Asian data

librarians.

Lastly, UI convenience of the EADL website should be improved. One of the purposes of the pilot service operation would be upgrading UI to be suitable for actual user usage pattern.

5. Conclusion

In this study, the contents and services of EADL were analyzed and the strengths and weaknesses were identified to derive improvement points. Since starting pilot service on December 17 2020, EADL has been operating the service for over a month so far. In the future, we need to collect more materials, attract participation from various institutions and reflect the needs of real users. Through these efforts, EADL should be turned into a useful digital library service. EADL is a very meaningful attempt in that it provides digitized materials from Korea and Japan to anyone online. EADL should also collect rare materials related to East Asia that are scattered around the world and provide them to related researchers. Research and planning should be continued to make EADL a useful digital library service for users

References

East Asia Digital Library. http://eadl.asia/.
Europeana. http://www.europeana.eu/en.
Fukubayashi, Yasuhiro et al, (2020). East Asia Digital Library: An example of an international joint project in East Asia. デジタルアーカイブ学会誌, 4(S1), S80-3.
National Diet Library. http://www.ndl.go.jp/en/index.html.
National Library of Korea. http://www.nl.go.kr/.
NLK LOD. http://lod.nl.go.kr/home/
Park, JinHo (2020). East Asia Digital Library: EADL. Seoul: National Library of Korea.
Wikipedia. "Linked Data." http://en.wikipedia.org/wiki/Linked-data.
World Digital Library. http://www.wdl.org/en/.