FROM THE PRESIDENT

PEER-REVIEWED ARTICLES

An Alternative Method of Title Authority Control: The Shifted-Authority Control Model for Chinese Classics
Maiko Kimura

In Search of Hong Kong Literature in a Digital Hub: An Evaluation of Hong Kong Literature Database from the Perspective of Researchers
Chui Yu Cheung, Dickson K.W. Chiu and King Kwan Li

REPORTS

Charlene Chou

IN MEMORIAM

INSTITUTIONAL NEWS

NEW APPOINTMENTS

PUBLICATION NEWS
From the President

As the autumn is turning the colors to red and gold, our campuses are also wakening up to the hustling footsteps and energy of new and returning students and faculty. Many have just come back from their travels for business, family, or leisure. I’d like to take this opportunity to wish that all CEAL colleagues have had a productive and re-energizing summer and will have a fruitful and rewarding 2019-2020 academic year.

Since the last CEAL annual meeting in March, the Executive Board (EB) and task force members have made significant progress on CEAL initiatives and in planning for the 2020 annual meeting. As you have heard from our vice president, Hana Kim, CEAL will present its first Distinguished Service Award at the 2020 annual meeting and the task force is calling for nominations. I invite CEAL members to participate actively and submit your nominations. As an organization, this award raises the flag for the values and mission of CEAL. The recipient will represent the highest quality and achievements of CEAL members.

The Task Force on CEAL Leadership Institute 2020 chaired by CEAL Membership Committee Chair Xiang Li has been working hard to organize the first in a long-time CEAL leadership training program. A survey has been distributed to CEAL members to collect ideas and suggestions for training topics. We are honored to have Maureen Sullivan, a past president of ALA and renowned educator in the library profession, to be our lead instructor. She has been working with the CEAL task force on planning and designing the training curriculum. The leadership institute will include a preconference of CEAL annual meeting at Harvard University on March 17 and three to four webinars all catered to meet the particular needs of CEAL members. In addition, the CEAL annual meeting will have a plenary session with outstanding library leaders as keynote speakers. The task force is also working with grant organizations and donors to raise funding to support the leadership institute.

The CEAL Organizational Handbook Task Force led by CEAL Secretary Yao Chen has made great progress putting together the first edition of the handbook. I thank all the past and present EB members for contributing to the entries of this important document, which will serve as a very useful tool to ensure continuity and development of our organization. This is a detailed and demanding task. I commend the leadership of Yao and her team for their dedication and encourage all CEAL officers to continue contributing to the handbook to make the document as complete as possible.
In July, CEAL submitted its annual report to AAS. We reported on the successful completion of the initiative by the Task Force on the 2019 CEAL Annual Program: Panel of EA Librarians, the progress of the other CEAL initiatives, and the significant achievements of CEAL committees, task forces, and affiliate groups. CEAL Treasurer submitted a detailed report on the status of CEAL finance. Our report has received very positive comments from AAS President and Executive Director. The report is now available to CEAL member upon request to CEAL Secretary.

CEAL EB has also been working on the 2020 CEAL annual program, which will continue the theme of “Strengthening our organization and empowering CEAL members to meet the challenges of the digital age.” Thanks to your responses to the 2019 post-conference survey, our programs in 2020 will focus on bringing top-notch speakers. The topics of the two plenary sessions in 2020 will be: challenges and opportunities of digital scholarship for scholars and librarians; and building leadership competencies for East Asian librarians to succeed in the digital age. CEAL committee chairs are also proactively planning for committee sessions and preconference. We look forward to presenting a rich repertoire of activities to CEAL member in Boston next March.

Last but not the least, I’d like to thank our JEAL Editor, Charles Fosselman, for his excellent and dedicated work to have put together another content-rich issue of JEAL. There are two peer-reviewed research articles in this issue discussing topics directly from our field of East Asian librarianship. Enjoy the reading!

Zhijia Shen
CEAL President
An Alternative Method of Title Authority Control: 
The Shifted-Authority Control Model for Chinese Classics

Maiko Kimura
School of Library and Information Science, Keio University

Background

Catalogs and authority control for Chinese classics

The Functional Requirements for Bibliographic Records (FRBR) and the IFLA Library Reference Model (IFLA LRM) (International Federation of Library Associations and Institutions 2017), which consolidated FRBR and its related conceptual models Functional Requirements for Authority Data (FRAD) and Functional Requirements for Subject Authority Data (FRSAD), have proposed a new model for library-cataloging. In this model, several related bibliographic records are connected and presented together in groups. Works, the top-level concept of the hierarchy of the bibliographic data, are presented by including groups of several materials with the same content but different titles and/or related materials. This model forms the basis of Resource Description and Access (RDA) (Canadian Library Association and American Library Association, Joint Steering Committee for Development of RDA 2017), a cataloging code adopted by many libraries worldwide.

Applying the work-expression-manifestation-item (WEMI) hierarchy shown in FRBR and IFLA LRM can also be suitable for cataloging rare materials or classic works; Attig (2003) stated that although catalogers of rare materials must generally be flexible in order to address the special characteristics of such materials and their uses, standards and guidelines are still important, as they provide benefits regarding the global sharing of bibliographic information. However, Fell and Lapka (2016) highlighted that “the FRBR assessment does not accurately account for the importance of a number of data elements for the description of rare materials,” while Martín and Prada (2016) noted insufficiencies in RDA in terms of instructions for describing special features of rare materials.

Special measures are required for cataloging rare materials, as they have special characteristics. Therefore, since 1981, sets of instructions concerning cataloging standards for rare materials have been advocated by the Bibliographic Standards Committee, Rare Book and Manuscript Section (RBMS), of the Association of College and Research Libraries (ACRL) and the Library of Congress. Further, a series of manuals
entitled *Descriptive Cataloging of Rare Materials* (DCRM), which relate to several types of rare materials, were published between 2007 and 2016. After several years of efforts, an RDA-compatible set of guidelines for cataloging rare materials, entitled *RBMS Policy Statement to RDA* (RBMS PS), was submitted by the DCRM Task Force of the ACRL/RBMS at the 2017 American Library Association Annual Meeting (American Library Association 2017). Thus, RBMS PS is tasked with addressing the abovementioned issues raised by previous researchers concerning cataloging rare materials. Meanwhile, regarding rare Chinese books, which mainly relate to Chinese works published before 1796 (i.e., by the end of the Qianlong era (1736-1795) of the Qing Dynasty, covering a part of the Chinese classics published before 1912), the Council on East Asian Libraries’ Committee on Chinese Materials has been assigned the task of revising the *Cataloging Guidelines for Creating Chinese Rare Book Records in Machine-Readable Form* (published in 2000 and revised in 2009), to ensure that RDA can be applied to rare Chinese materials (Council on East Asian Libraries, 2017). Thus, cataloging and bibliographic control, including for Chinese classics, is clearly trending toward the use of FRBR.

Several projects have applied FRBR or RDA to catalogs of rare materials. Katic (2001) tested FRBR by using this style to present bibliographic records for rare books in Latin and their German and Italian translations, concluding that structuring records using the WEMI hierarchy is sufficient for meeting users’ needs. Meanwhile, RDA was implemented in a project concerning rare materials that was undertaken by researchers from the University of Kent (Caplehorne and Dickinson, 2015), and also in a project conducted by 16 organizations from German-speaking countries (Aliverti and Behrens, 2016). Further, MacDonald and Lawrence (2016) stated that the RDA approaches for collected works, translations, and works issued together are potentially suitable for the cataloging of rare materials. Demonstrating this, Tokita et al. (2012) investigated 158 major Japanese classics and, as each work featured several expressions and a large number of manifestations, they concluded that applying the FRBR model to cataloging such classics is beneficial.

Similar to Japanese classics, there can be many different expressions and manifestations of a work in Chinese classics, each possessing a different title (Downing, 2014). Moreover, it is common for an item to have different titles in different areas (e.g., first page of the main text, center column, and inner cover). Therefore, to meet the demands of users who wish to retrieve all items relating to an intended work, without omission, several access points must be provided in the relevant title authority record; namely, a title authority record must include the authorized title and variant titles of the work.
Utilizing title authority control, a part of the WEMI hierarchy can be applied to the field of Chinese classics. Each bibliographic record is a surrogate of a Chinese classic item, which is “somehow derived from a common progenitor” (Smiraglia and Leazer, 1999), and an authorized title of the work to which the item belongs is also recorded. Further, various bibliographic records relating to the same work (i.e., the progenitor) can be grouped under the authorized title of the work. Even if there are no records of expressions or manifestations, users can at least find Chinese classics at the work-level. An example of the correspondence relationship between the WEMI hierarchy and a Chinese classic is shown in Figure 1. This figure is an example to help readers easily understand the idea of this study. Numerous exceptions must exist, especially in the expression- and manifestation-levels. A strict way to determine these levels for Chinese classics has not been developed yet and should be investigated in future studies.

![Image of WEMI hierarchy](https://example.com/image1)

**Figure 1.** An example of a Chinese classic with WEMI hierarchy.

### Current situation of title authority control in Japan and China

Few title authority records exist in Japan and China. The National Diet Library (NDL) in Japan, which possessed 4,678 title authority records (not limited to classical works) as of 1

---

March 2017, only conducts title authority control for Japanese classics (National Diet Library, n.d.; 2012a). Further, NACSIS-CAT, an online shared-cataloging system for academic libraries in Japan, operated by the National Institute of Informatics (NII), had 36,078 title authority records (not limited to classical works) as of March 2017 (National Institute of Informatics, 2017). Although Chinese classics are included in the NACSIS-CAT’s title authority files, it is assumed that there are few authority records for Chinese classics, since conducting authority control and creating authority records are not mandatory for member libraries (Miyazawa and Koto, 2012). Further, NACSIS-CAT cataloging guidelines of Japanese and Chinese classics (NII, 2011) have not yet established any rule to construct authorized access points for Chinese classics.

In October 2017, the National Library of China had approximately 60,000 title authority records (Niu, 2017). Further, according to the author's personal correspondence, in January 2018, the China Academic Library and Information System (CALIS), which is an academic library consortium funded primarily by the Chinese government, possessed 51,079 Chinese title authority records (mainly concerning series titles). Moreover, as part of a project by the Joint University Librarians Advisory Committee, which is a consortium of eight academic libraries in Hong Kong, the Hong Kong Chinese Authority Name Project operates a union database for Chinese records (Joint University Librarians Advisory Committee, n.d.), and in November 2016, this database held over 54,000 title authority records (Xu, 2016). Finally, according to an interview conducted by the author, in July 2013, the National Central Library in Taiwan had 18,252 title authority records (the abovementioned figures are not limited to classical works).

Since 2004, CALIS has especially maintained a database for Chinese classics, which also partially includes images of each item (CALIS, n.d.). However, it seems that the database is not authority-controlled.

Compared to the Library of Congress, which had 526,623 title authority records on June 1, 2013 (Virtual International Authority File 2018), the numbers of title authority records in the abovementioned libraries in Japan and China are very low. Unfortunately, it is impossible to know the exact number of title authority records for Chinese classics in these databases because there is no identifier on the authority records proving that the record is for a Chinese classic.

Instead of title authority control, book catalogs of Chinese classics, in which the works are organized in a special classified order, are used. In addition, a reference book entitled Tong shu yi ming hui lu [The list of different titles for the same book] (Du and Wang
which is designed to facilitate the checking of different titles concerning Chinese classics, has been revised several times and is widely used. However, without title authority data, it is impossible to either conduct comprehensive title authority control for Chinese classics or apply the WEMI hierarchy to such works.

**Features of Chinese classics**

In contrast to modern Chinese materials, Chinese classics have four distinguishing features:

1. *They do not increase.* Although there are a few exceptions (Center for Informatics in East Asian Studies 2009), Chinese classics have primarily been defined as works written by Chinese people in the Chinese language before approximately 1912 (Standardization Administration of the People’s Republic of China, 2008; Japan Library Association, 2006). This is because the Republic of China was established after the Qing dynasty ended in 1911. Therefore, unless new works are discovered, the number of Chinese classics cannot increase anymore.

   There are no statistics on how many Chinese classics exist in Japan. According to Hao (2014), China has about 50 million physical volumes of Chinese classics. Wu (2010) stated that there are about 30 million physical volumes. The Chinese government has implemented the “Chinese ancient books census” project and the exact figure will be known in 2020 (Ministry of Culture of the People's Republic of China 2017).

2. *Each book is different from others on the item-level.* Chinese classics are, and should be, handled as individual items of a manifestation, with each item being unique, even though they have been printed using the same printing woodblocks. A set of printing woodblocks for a manifestation could be preserved and used many times. With time, a part of a woodblock or even a part of the set woodblocks might become broken, deleted, re-carved, corrected, or altered by another woodblock. Such differences exist among items printed around the same time, implying that each item may have unique texts. In addition, each item has its own item-specific carrier characteristics, such as handwritten marginalia, seals, and binding information. Such information is extremely important and might be the subject of East-Asian studies. Therefore, a bibliographic record has been created for each item in Japan (Japan Library Association (JLA) 2006). Although two items may resemble each other, and may even have no item-specific characteristics, their bibliographic records are maintained separately. This is a common feature among rare materials from other cultural spheres (Attig, 2003).
3. **Chinese classics were also produced in Japan.** Chinese classics, both content-wise and physically, are widespread across the entire cultural sphere of Chinese characters: China, Japan, Korea, and Vietnam. In Japan, for example, many Chinese classics were imported and republished during the Edo period (1603–1867) (Zhou, 2012). As a result, Chinese classics in Japan include not only items imported from China, but also a massive amount of items republished in Japan.

4. **Including numerous variant Chinese characters.** Even in Mainland China and Japan, where people currently do not use traditional Chinese characters in their daily lives, bibliographic data for Chinese classics is customarily recorded in traditional Chinese characters (SAC 2008; CIEAS 2009). This is because Chinese classics must be written in traditional Chinese characters. However, in some cataloging systems, variant characters, such as Japanese Kanji, remain in some records. For example, the Universal multiple-octet coded Character Set (UCS) was adopted by NACSIS-CAT in 2000. Before this, when a cataloger could not input a Chinese character to the system because the character set did not cover them, the cataloger used corresponding Japanese Kanji characters instead (Yonezawa, 2001). Such use of Japanese Kanji characters may cause retrieval omissions.

In addition, variant characters exist among traditional Chinese characters themselves. For example, both "鑒(U+9452)" and "鍳(U+9373)" are traditional Chinese characters with the same meaning and pronunciation, but they are different characters (their Unicode code points (U+XXXX) are different). Chinese classics include numerous variant characters (Li, 2015). Authority control for titles of Chinese classics is very important because each of the several items belonging to the same work sometimes has almost the same title, with a variant Chinese character. Without authority control, this also may cause retrieval omissions.

To avoid such retrieval omissions, the Chinese-Japanese-Korean (CJK) unified index has been introduced to NACSIS-CAT (Miyazawa, 2006). Applying the index, the system automatically broadens the search target and variant characters can also be retrieved. As Miyazawa (2006) mentioned, although it may cause inappropriate results (i.e., noises) because the index tends to unify as many variant characters as possible, the index is valuable because users do not need to worry about variant characters. A similar index is also applied to the Zenkoku Kanseki Database, a dominant database used for retrieving Chinese classics in Japan (Yasuoka, 2002). Currently, at least in Japan, each retrieval system applies its own index with different ranges of variant characters (Nagasaka, 2014). The NDL Search is an integrated search engine of the catalogs and digital resources of Japanese public libraries, the National Diet Library, and other academic institutes (National Diet Library, 2012b). Nagasaka (2014) also pointed out that
the NDL Search only has an index for a small number of variant characters. In addition, since the number of variant Chinese characters is uncountable, such indexes cannot be perfect. There is always the fear of omission or noise.

5. Having difficulties of identification in manifestation-level. Items printed from one set of printing woodblocks may be assumed to belong to one manifestation. However, identifying a manifestation is very difficult without sophisticated expertise. For example, a set of newly-carved printing woodblocks are simulant of an older set of woodblocks and sometimes bear a striking resemblance to the older set (Zhou, 2016). In such a case, it is not easy to determine whether two sets belong to the same manifestation or not. Further, it is very difficult to decide whether a part of an item printed from altered woodblocks should be considered as a part of the original manifestation, or if they should be assumed as a new manifestation altogether. Therefore, at least in Japan, it is commonly understood that identification of manifestation is a professional task that should be performed by experts, such as researchers and bibliographers, and not by library catalogers.

Taking the above into consideration, the author proposes a new authority control model for Chinese classics in Japan. Usually, authority control requires links between authority records and bibliographic records. However, even if we could successfully create comprehensive authority data for the titles of Chinese classics, linking them to each bibliographic record retrospectively would be cumbersome. Further, adding the access points for all existing bibliographic records is also nearly impossible without a large labor force. Thus, the proposed model, named the shifted-authority control (SAC) model, pursues improved recall with less expense, allowing users to find more relevant items than ever. The SAC can improve recall more effectively than the current situation in Japan, as it utilizes title authority data, and has less expense than a fully authority-controlled model, as the SAC does not require links between authority and bibliographic records.

Specifically, the purpose of this study is to investigate the effectiveness of the SAC model. Although this study limits its scope to applying the SAC model to bibliographic records in Japan, the model itself and title authority data that the author prepares can also be applied to bibliographic databases of Chinese classics in other regions.

Methods

SAC Model

Figure 2 shows the current authority control situation in Japan. Under the current situation in Japan, although the bibliographic record ID 54321 is the relevant record, the user cannot find it because it does not include the search word and the title is not
authority-controlled. The title “九章重差” (Jiu zhang chong cha) is a search word used by the user in Figure 2. The work known under the title “九章重差” (Jiu zhang chong cha) is also known under a variant title “海島算經” (Hai dao suan jing). The two titles are completely different in their Chinese characters, pronunciations, and meanings.

On the other hand, in the normal authority control model shown in Figure 3, a bibliographic record has an authorized access point (AAP) for the work, which is linked to the work’s corresponding authority record. The AAP is assigned to only one authorized title among several different titles for a work. In the authority record, the AAP is recorded. Further, the work’s variant access points (VAPs), which are variant titles that are not selected as the AAP, are also recorded along with the authors’ names, sources of information, etc. In this model, even when a user searches for a title that does not appear in the bibliographic records, the system can suggest it or automatically retrieve an appropriate authority record from the authority database. Through the link between the authority and bibliographic records, the user can obtain all relevant bibliographical records which have the AAP for the work, including the bibliographic record ID 54321 in Figure 3. The normal model (Figure 3) needs the AAP for the work to be added to each existing bibliographic record with links to the work authority records, as well as the creation of authority records. Although some systems work out authority control without links between bibliographic and authority records, AAPs controlled by authority records must be added to bibliographic records. Since titles for Chinese classics had not been controlled in Japan, achieving the normal model requires an enormous amount of labor; this seems almost impossible in Japan.
The SAC model proposed in this study is shown as model A in Figure 4. For comparison, the author also set a model B, the authority records of which are easier to construct, but the omissions in its retrieval results seem to improve.

Model B, named as the independent authority records model, is a more realistic approach for work title authority control for Chinese classics in Japan. Although the SAC
model (model A) needs the construction of authority records, AAPs are not assigned to every bibliographic record. When a user searches for a title, the system searches the authority database first to find the work’s AAP and VAP. Then, the system automatically searches the bibliographic database using the AAP and the VAP as search words. In case of Figure 4, the final search results obtained by the user include bibliographic record ID 54321. However, the results may also include some records that are not relevant to the search query. This is because bibliographic records do not have controlled AAP and the retrieval has no choice but to run a keyword search of the entire bibliographic database.

Besides the increasing noise, an anticipated problem with model B is retrieval omissions. Chinese classics may contain variant Chinese characters in their titles as explained in the fourth feature. In case a bibliographic record is recorded in variant Chinese characters that are different from both AAP and VAPs, the record may not be retrieved.

In the SAC model (model A), determining the link between bibliographic records and authority records is not required, and AAPs are not assigned to every bibliographic record, as well as model B. In this model, however, each title authority record contains the several item titles assigned to each work. When a user searches for a title, the system searches the authority database first to find the work’s AAP, VAPs, and item titles. Then, the system automatically searches the bibliographic database using those access points and item titles as search words.

Compared to model B, there are two advantages of model A. First, since various item titles and their variant titles are recorded in authority files, it can prevent retrieval omissions caused by variant Chinese characters, even for a system that does not implement an index for variant Chinese characters. Second, since the SAC model refers to the FRBR (now IFLA LRM) model and adopts work and item-level records, users can also directly search the authority database and obtain the representative item titles of each work. Although bibliographic records and authority records are not linked in both models A and B, model A’s approach can reveal relationships between the works and items. This is beneficial to users, including librarians, who may seek exact item titles belonging to a particular work.

It is assumed that in both models A and B, once the authority database is constructed, catalogers in libraries will not be required to create or maintain authority records again; this is because the number of Chinese classics cannot increase any further, as noted in the previous section.
Data construction

For the purpose of investigating the effectiveness of the SAC model, first, title authority data were constructed. Specifically, the data construction process was as follows:

1. Choosing an AAP for each work.

   The title authority data used were based on *Tong shu yi ming hui lu* (Du and Wang, 2000). Although *Tong shu yi ming hui lu* provides different titles for each Chinese classic, it does not specify whether titles should be AAPs or VAPs. Therefore, the author selected an AAP for each work, applying the following order of priority. (a) The title that returns the most search results in the Zenkoku Kanseki Database (CIEAS, n.d.) is preferred; (b) a title that is not used in other works is preferred; (c) a title that appropriately represents the subject of the work is preferred. To distinguish between different works sharing the same title, for some works the second priority takes precedence over the first; this removes the risk of AAP duplication.

2. Recording the AAP and VAPs for each title authority record.

   At this stage, a prototype of a title authority database for Chinese classic works, *KWMA-san* (Kimura, 2017), was constructed, and a title authority record was created for each work. Each record contained an AAP selected by the author, while non-preferred titles were recorded as VAPs. The works listed in *Tong shu yi ming hui lu* are written in simplified Chinese characters; however, the author entered AAPs, VAPs, authors’ names, and other information into the database using traditional Chinese characters.

   To obtain item titles for each work, CiNii Books was searched using the corresponding AAPs and VAPs. Then, appropriate bibliographic records were selected and imported into *KWMA-san* through the OpenSearch API of CiNii Books (NII, 2015). This suggests that each work possessed representative item titles. When CiNii Books returned many records, those with titles that differed from other records were selected; however, when CiNii Books did not return results, the item title could not be recorded. Each authority record had between one and 22 unique item titles, with the average being five.

   With regard to the third step, the reasons for choosing CiNii Books as a source of item titles are as follows:
(a) Although there is no authoritative cataloging code for Chinese classics in Japan, usually only one title is chosen and recorded in such catalogs, even if there are various titles for an item. However, because of the cataloging guidelines for Japanese and Chinese classics of NACSIS-CAT (NII, 2011), which is a union-catalog database system for academic libraries in Japan and the source of bibliographic records for CiNii Books, bibliographic records for Chinese classics tend to have various titles in variant title fields. Therefore, it was easy to obtain several item titles from CiNii Books records.

(b) The Zenkoku Kanseki Database usually does not provide variant titles. In addition, it does not provide an API that would facilitate the utilization of its bibliographic records as sources.

(c) NDL Search integrates bibliographic records of Japanese public libraries, the National Diet Library, and other academic institutes. Bibliographic records from public libraries, however, are usually very simple and not suitable for obtaining item information.

(d) The author could not find any other dominant bibliographic database for retrieving Chinese classics in Japan.

_Tong shu yi ming hui lu_ has over 13,500 titles (Du and Wang 2000), listed in the order of the stroke counts of the first Chinese character in each title. In this list, each work has at least two titles, and some have more than five; the author has not yet calculated the total number of featured works. Data imputing began with the first page of _Tong shu yi ming hui lu_ and, as of January 2018, 270 title authority records have been inputted into KWMA-san.

The models between the WEMI hierarchy and the SAC's are shown in Figure 5. As IFLA LRM and FRBR are conceptual models, the author merely adopts work and item entities for the data model of the SAC. While the author also assumes that the manifestation entity is useful for identifying Chinese classics, manifestation has not been adopted in this implementation because 1) bibliographic records for Chinese classics are created at the item-level, and not in the manifestation-level in Japan, 2) manifestations of Chinese classics are too difficult to identify for librarians in Japan. Figure 6 presents an example of a title authority record in KWMA-san.
Retrieval experiments

In this study, the author substituted three types of search queries for models A, B, and the current situation in Japan (Figure 2). As model A in Figure 4 shows, in the SAC approach,
when the user searches the bibliographic database, the system automatically retrieves the authority database first and obtains the AAP, VAPs, and item titles of the work needed. Then, the system returns to the bibliographic database and searches it with those access points. Instead of constructing a retrieval system for the SAC model, the author runs a Boolean OR search for the AAP, VAPs, and all item titles (query iii) for 80 works.

CiNii Books, NDL Search, and the Zenkoku Kanseki Database were searched using the following query types: (i) AAP only; (ii) a Boolean OR search for the AAP and all VAPs. Query i represents the current situation in Japan (Figure 2) and query ii represents model B in Figure 4. Two queries are used for comparison with query iii.

Retrievals were conducted for 80 works, which were selected in the order they appear in Tong shu yi ming hui lu. However, works that had no item title (i.e., those that did not return any results from CiNii Books during the third step of the data construction process) were not included in those 80 works, as it was supposed that search queries ii and iii would return identical findings. Similarly, if all item titles for a work corresponded to the work's AAP or VAPs, the work was also excluded from this experiment, because both query ii and iii would again return identical results.

After each retrieval, the author examined whether individual bibliographic records in the search results for the given work could be considered items of the work. In other words, if an item represented in a bibliographic record belonged to, or was derived from, the progenitor “work” that intended to retrieve, the bibliographic record would be identified as a correct record. When a bibliographic record represents an item that corresponds to various works, the record is judged to be the correct record for each of these works, under the condition that the title representing each work appears in the title, variant title, or note field of the bibliographic record. The correct bibliographic records identified for each work through this process are assumed to be linked to each authority record in the normal authority control model (Figure 3).

By comparing the search results of search query i and ii, the effectiveness of the SAC model could be investigated. Search query i represents a situation in which there is no title authority control, and where a user has searched for a work using only one already-known title. Meanwhile, query ii represents model B (the independent authority records model) in Figure 4.

The AAPs and VAPs presented in traditional Chinese characters were used for retrieval. Item titles, on the other hand, can be recorded in various types of characters in KWMA-san, not limited to traditional Chinese characters because NACSIS-CAT allows all
types of Chinese characters (including Japanese Kanji) for their records. Therefore, item titles searched using query iii were inputted as they appear in the bibliographic records in CiNii Books.

After identification, recall, precision, and F-measure, which is a weighted harmonic mean of recall and precision, were calculated based on the correct records, as follows:

\[
\text{Recall} = \frac{\text{the number of correct records}}{\text{the number of correct records obtained through query iii}}
\]

\[
\text{Precision} = \frac{\text{the number of correct records}}{\text{the number of records retrieved}}
\]

\[
F - \text{measure} = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}}
\]

The results obtained by applying query iii were assumed to include the largest number of correct records for the searches in question. Therefore, the recall of query iii was always 100%, and the recall of queries i and ii were determined by comparing their number of correct recalls to those of query iii. It was also assumed that the recall of query iii (100%) is the same as the recall of model A.

Ideally, for each bibliographic database queries i to iii should be identical; however, the Zenkoku Kanseki Database does not allow Arabic numbers, colons, semicolons, and periods in search queries. Therefore, when applying query iii, Arabic numbers that appeared in item titles were changed to numerical figures written in Chinese characters. This change had no practical impact on the search results of CiNii Books and NDL Search since, in both databases, there was no change in the number of retrieved results. Further, for the Zenkoku Kanseki Database, colons in item titles were also deleted in search queries, while semicolons and periods were simply changed to the “+” character, which means Boolean OR search. Additionally, NDL Search does not allow spaces in search
queries, except on both sides of Boolean operators; therefore, spaces were deleted in queries for NDL Search.

Results and Discussion

Summary of retrievals and identifications

Each bibliographic database was searched for each work using the three query types, and then correct records were identified by the author. If any of the three queries returned no records or none that were identified as correct, the recall or the precision of the query in question could not be calculated. Consequently, in such cases all results for the work in question were excluded from the overall results. Therefore, although all 80 works were searched, the numbers of works (n) shown in Tables 1 and 2 differ depending on the bibliographic databases.

As evident in Table 1, the number of retrieved/correct records differs for each work. For example, in CiNii Books, one work returned 1,485 bibliographic records, of which 1,314 were identified as correct. On the other hand, 10 works returned only one bibliographic record, with each being correct.
Table 1. Summary of retrievals and identifications.

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Query</th>
<th>M</th>
<th>SD</th>
<th>Mdn</th>
<th>MAD</th>
<th>MIN</th>
<th>MAX</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>i</td>
<td>Records Retrieved</td>
<td>40.10</td>
<td>172.37</td>
<td>8</td>
<td>8.90</td>
<td>1</td>
<td>1485</td>
<td>7.65</td>
<td>60.50</td>
<td>19.52</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Correct Records</td>
<td>34.35</td>
<td>152.18</td>
<td>7</td>
<td>7.41</td>
<td>1</td>
<td>1314</td>
<td>7.71</td>
<td>61.28</td>
<td>17.23</td>
</tr>
<tr>
<td></td>
<td>ii</td>
<td>Records Retrieved</td>
<td>104.74</td>
<td>330.56</td>
<td>13</td>
<td>14.08</td>
<td>1</td>
<td>1782</td>
<td>4.30</td>
<td>17.94</td>
<td>37.43</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Correct Records</td>
<td>43.36</td>
<td>158.65</td>
<td>9</td>
<td>10.38</td>
<td>1</td>
<td>1314</td>
<td>6.84</td>
<td>50.31</td>
<td>17.96</td>
</tr>
<tr>
<td></td>
<td>iii</td>
<td>Records Retrieved</td>
<td>172.01</td>
<td>446.47</td>
<td>18</td>
<td>21.50</td>
<td>1</td>
<td>2220</td>
<td>3.53</td>
<td>11.96</td>
<td>50.55</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Correct Records</td>
<td>45.58</td>
<td>169.63</td>
<td>10</td>
<td>10.38</td>
<td>1</td>
<td>1414</td>
<td>6.95</td>
<td>51.77</td>
<td>19.21</td>
</tr>
<tr>
<td></td>
<td>i</td>
<td>Records Retrieved</td>
<td>41.40</td>
<td>200.33</td>
<td>3</td>
<td>2.97</td>
<td>1</td>
<td>1482</td>
<td>6.76</td>
<td>45.63</td>
<td>27.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Correct Records</td>
<td>26.93</td>
<td>113.50</td>
<td>3</td>
<td>2.97</td>
<td>1</td>
<td>830</td>
<td>6.46</td>
<td>42.60</td>
<td>15.30</td>
</tr>
<tr>
<td></td>
<td>ii</td>
<td>Records Retrieved</td>
<td>65.58</td>
<td>214.97</td>
<td>5</td>
<td>5.93</td>
<td>1</td>
<td>1482</td>
<td>5.40</td>
<td>31.68</td>
<td>28.99</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Correct Records</td>
<td>33.87</td>
<td>122.88</td>
<td>4</td>
<td>4.45</td>
<td>1</td>
<td>830</td>
<td>5.30</td>
<td>29.85</td>
<td>16.57</td>
</tr>
<tr>
<td></td>
<td>iii</td>
<td>Records Retrieved</td>
<td>121.80</td>
<td>293.18</td>
<td>9</td>
<td>11.86</td>
<td>1</td>
<td>1483</td>
<td>3.34</td>
<td>11.15</td>
<td>39.53</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Correct Records</td>
<td>33.96</td>
<td>122.85</td>
<td>4</td>
<td>4.45</td>
<td>1</td>
<td>830</td>
<td>5.30</td>
<td>29.86</td>
<td>16.57</td>
</tr>
<tr>
<td></td>
<td>i</td>
<td>Records Retrieved</td>
<td>35.52</td>
<td>46.84</td>
<td>21</td>
<td>26.69</td>
<td>1</td>
<td>267</td>
<td>2.65</td>
<td>8.50</td>
<td>5.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Correct Records</td>
<td>33.25</td>
<td>45.42</td>
<td>18</td>
<td>22.24</td>
<td>1</td>
<td>267</td>
<td>2.91</td>
<td>10.20</td>
<td>5.32</td>
</tr>
<tr>
<td></td>
<td>ii</td>
<td>Records Retrieved</td>
<td>74.14</td>
<td>161.47</td>
<td>31</td>
<td>35.58</td>
<td>2</td>
<td>855</td>
<td>3.90</td>
<td>14.95</td>
<td>18.90</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Correct Records</td>
<td>42.21</td>
<td>58.27</td>
<td>24</td>
<td>29.65</td>
<td>2</td>
<td>334</td>
<td>2.92</td>
<td>9.94</td>
<td>6.82</td>
</tr>
<tr>
<td></td>
<td>iii</td>
<td>Records Retrieved</td>
<td>22.89</td>
<td>467.42</td>
<td>53</td>
<td>68.20</td>
<td>2</td>
<td>2844</td>
<td>3.70</td>
<td>15.38</td>
<td>54.71</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Correct Records</td>
<td>43.11</td>
<td>58.18</td>
<td>25</td>
<td>29.65</td>
<td>2</td>
<td>335</td>
<td>2.93</td>
<td>10.02</td>
<td>6.81</td>
</tr>
</tbody>
</table>
**Statistical analysis**

Table 2 presents the means for recall, precision, and F-measure, which were calculated from search results and identifications. In addition, a repeated one-way analysis of variance (ANOVA) was used to compare the mean values between the three query types.

**Table 2.** The means for recall, precision, and F-measure regarding the search results.

<table>
<thead>
<tr>
<th>Bibliographic database</th>
<th>CiNii Books (n = 78)</th>
<th>NDL Search (n = 55)</th>
<th>Zenkoku Kanseki DB (n = 73)</th>
</tr>
</thead>
<tbody>
<tr>
<td>query</td>
<td>query</td>
<td>query</td>
<td>query</td>
</tr>
<tr>
<td>query</td>
<td>query</td>
<td>query</td>
<td>query</td>
</tr>
<tr>
<td>query</td>
<td>query</td>
<td>query</td>
<td>query</td>
</tr>
<tr>
<td>M</td>
<td>0.764</td>
<td>0.945</td>
<td>1.000</td>
</tr>
<tr>
<td>F(2, 154) = 26</td>
<td>F(2, 108) = 20.5</td>
<td>F(2, 144) = 39.6</td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>η²</td>
<td>0.338</td>
<td>0.179</td>
<td>0.225</td>
</tr>
<tr>
<td>1-β (α = 0.10%)</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>M</td>
<td>0.922</td>
<td>0.787</td>
<td>0.667</td>
</tr>
<tr>
<td>F(2, 154) = 26</td>
<td>F(2, 108) = 15.7</td>
<td>F(2, 144) = 52.5</td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>η²</td>
<td>0.125</td>
<td>0.092</td>
<td>0.266</td>
</tr>
<tr>
<td>1-β (α = 0.10%)</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>M</td>
<td>0.793</td>
<td>0.814</td>
<td>0.727</td>
</tr>
<tr>
<td>F(2, 154) = 3.4</td>
<td>F(2, 108) = 5.5</td>
<td>F(2, 144) = 18.5</td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>0.036</td>
<td>0.005</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>η²</td>
<td>0.026</td>
<td>0.037</td>
<td>0.117</td>
</tr>
<tr>
<td>1-β (α = 0.10%)</td>
<td>0.36</td>
<td>0.90</td>
<td>1</td>
</tr>
</tbody>
</table>
The ANOVA showed that there was a strong significant difference between the mean values for recall, precision, and F-measure in all bibliographic databases, except for the F-measures of CiNii Books and NDL Search (p < 0.05 and p < 0.01, respectively). Measuring the strength of a relationship ($\eta^2$) is one of the measures of effect size facilitated by an ANOVA; according to Cohen (1988), effect-size magnitudes are small, medium, and large when $\eta^2 > 0.01$, $\eta^2 > 0.06$, and $\eta^2 > 0.14$, respectively. Consequently, it was found that the F-measures of CiNii Books and NDL Search had relatively small magnitudes.

Further, when statistical power (1-$\beta$) is less than 0.8, the probability of making a Type II error (i.e., obtaining a false negative) exists (Cohen 1992). It was consequently determined that the F-measure for CiNii Books contained this risk.

Table 3. Results of pairwise comparisons (Holm’s method)

<table>
<thead>
<tr>
<th></th>
<th>CiNii Books</th>
<th>NDL Search</th>
<th>Zenkoku Kanseki DB</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>query ii</td>
<td>query iii</td>
<td>query ii</td>
</tr>
<tr>
<td>Recall</td>
<td>query i</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>query ii</td>
<td>-</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>query iii</td>
<td>-</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Precision</td>
<td>query i</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>query ii</td>
<td>-</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>query iii</td>
<td>-</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>F-measure</td>
<td>query i</td>
<td>$0.491_a$</td>
<td>$0.243_a$</td>
</tr>
<tr>
<td></td>
<td>query ii</td>
<td>-</td>
<td>0.02</td>
</tr>
</tbody>
</table>

a Not statistically significant

Next, Holm’s sequential Bonferroni procedure for pairwise comparisons was applied. As Table 3 shows, this procedure also returned a high number of significant differences (p < 0.001) between most of the queries in regard to recall and precision. However, there was no statistical difference between the F-measures of queries $i$ and $ii$ (p = 0.49) and $i$ and $iii$ (p = 0.24) in CiNii Books, queries $i$ and $ii$ (p = 0.50) in NDL Search, and queries $i$ and $ii$ (p = 0.12) in the Zenkoku Kanseki Database.
Factors affecting the results

Since recalls between queries $i$ and $iii$ and queries $ii$ and $iii$ had significant differences, it could be concluded that the proposed SAC model, which is represented by query $iii$, increases the number of results returned. On the other hand, it also lacks a certain degree of precision. For some query pairs, the differences concerning the F-measures were not significant.

There were more item titles with a semicolon or a period (i.e., items without collective titles) than the author expected (e.g., the second item title (Item id: 610) in Figure 6). In fact, of the 80 works searched, 32 had colons or periods in their item titles. Since the Zenkoku Kanseki Database does not allow semicolons or periods in search queries, the individual titles in these items were Boolean OR searched. This is considered to have decreased the precision of the search results. Further, Welch’s t-tests revealed that for query $iii$, the difference between the precision rates for searches of 73 general works and of 42 works that did not feature items without collective titles was statistically significant ($t(91) = 2.588, p = 0.011$); in fact, among the latter works the mean precision of query $iii$ in terms of the Zenkoku Kanseki Database increased to 0.763.

In some bibliographic records in CiNii Books for items without collective titles, the second and subsequent titles were recorded in variant title fields. Such a practice is unnecessary, although the cataloging guidelines of NACSIS-CAT do not prohibit it. An item without a collective title usually corresponds to various works that do not share a common progenitor, implying that because of the inclusion of CiNii Books’ records, the authority records in KWMA-san may have included irrelevant item titles. Naturally, such irrelevant item titles decrease the precision of searches of CiNii Books and NDL Search; however, eliminating such irrelevant titles from KWMA-san is impossible, as CiNii Books records do not have any designator to differentiate them from normal variant titles.

Comparing the recalls of queries $ii$ and $iii$ with those of query $i$ implies that the approaches of models A and B are always better than no authority control (Figure 2). Despite there being no link between bibliographic and authority records, the existence and utilization of an authority database can improve recall, and this may help users find more appropriate materials to use. The necessity of title authority control for Chinese classics in Japan should be re-evaluated and the current situation must be changed.

Conclusions

Table 4 shows the several aspects of authority control under the current situation in Japan (Figure 2), the normal authority control (Figure 3), the SAC model (model A in Figure 4), and the independent authority records model (model B in Figure 4). This study examined the efficiency of a new library-cataloging model, the SAC model. It was consequently found that this model could successfully obtain more results without requiring a cataloger to link between bibliographic records and authority records. Assuming that the normal authority control (Figure 3) can achieve nearly 100% recall without any noise, the precision of the SAC model is not preferable. However, the SAC model can avoid retrieval omissions as far as possible. This is highly important especially when a good index for variant Chinese
characters is not implemented as it expands the search results to the greatest extent possible. Further, the authority database with item titles of each work can provide work-item relationships to users, even without links between bibliographic and authority records.

Search results are dependent on bibliographic databases. As the Zenkoku Kanseki Database does not accept periods or semicolons, each part of the title of an item without a collective title had to be replaced with Boolean OR searched, consequently reducing searching precision. This suggests that the SAC model cannot be applied in an unrestricted manner; the search functions of each bibliographic database must be confirmed in advance. Unless the cataloging guidelines of NACSIS-CAT or the item title sources for KWMA-san change, increasing precision will be difficult. In other words, if tightly controlled bibliographic records are applied as item title sources, an increase in precision can be expected.

In RDA, the AAPs and VAPs representing works consist of (a) the AAPs for the agents (such as persons or corporate bodies) responsible for the works, if appropriate, and (b) the preferred titles of the works. This research, however, did not use agents’ AAPs. Contributors to Chinese classics (generally persons who were born before 1912) usually have several names, and those names are normally not authority-controlled in Japan. The difficulties in determining AAPs for Chinese personal names exceeds the scope of this research. Since bibliographic records for Chinese classics in Japan are currently based at item-level, collecting these items and forming manifestation- or expression-level records is also difficult. The author believes creating title authority records at work-level is the first essential step toward addressing these problems.

Taking into account that the precision of the SAC model is not optimal, the SAC model should be considered as merely an alternative to normal authority control and should be applied in situations where such authority control is insufficient. Although the scope of this research was limited to the field of Chinese classics, there is a possibility to apply the SAC approach to classics in other languages which share common features with Chinese classics. On the other hand, for modern materials, including reprints of Chinese classics that have modern covers and title pages, the author strongly recommends applying the normal authority control model.

Table 4. A comparison of authority control models.

<table>
<thead>
<tr>
<th>Features</th>
<th>Current</th>
<th>Normal</th>
<th>SAC</th>
<th>Independent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authority data</td>
<td>None</td>
<td>Need</td>
<td>Need</td>
<td>Need</td>
</tr>
<tr>
<td>Links between authority/bibliographic data</td>
<td>None</td>
<td>Need</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Item titles in authority data</td>
<td>None</td>
<td>None</td>
<td>Need</td>
<td>None</td>
</tr>
<tr>
<td>Recall</td>
<td>Low</td>
<td>High</td>
<td>High</td>
<td>Middle</td>
</tr>
<tr>
<td>Precision</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>Middle</td>
</tr>
</tbody>
</table>
Glossary

Chinese classics    Books created by Chinese people, mainly before 1912, in the Chinese language. They include printed books and handwritten manuscripts.

CiNii Books    A union catalog that enables the searching of bibliographic records in academic libraries in Japan. It can be accessed at: https://ci.nii.ac.jp/books/?l=en. Its data source is NACSIS-CAT, and its bibliographic records for Chinese classics are at the item-level.

Item-level record    Bibliographic record created for each item. In Japan, bibliographic records for Chinese (and Japanese) classics are created at the item-level. This means that manifestation-level records are not created. Instead, bibliographic data that should be recorded at the manifestation-level are recorded in the item-level records. An item may consist of multiple physical volumes. In this case, an item-level record is created for multiple volumes. When a volume is a compilation, it includes multiple works. Even so, usually an item-level record is created for the volume. Analytical titles for the volume can be, but not must be, recorded in the note field of the item-level record.

Item title    All titles and variant titles that appears in an item. Item titles for a classic may be equivalent to the titles for the manifestation, although some item titles, such as handwritten titles, may only appear in the item. Item titles do not include analytical titles.

NACSIS-CAT    An online shared-cataloging system for academic libraries in Japan that is operated by the National Institute of Informatics (NII). Each bibliographic record for Chinese classics is created at the item-level in NACSIS-CAT.

References


Du, Xinfu., and Wang, Jian. *Tong shu yi ming hui lu* (The list of different titles for the same book). Nanjing: Jiangsu gu ji chu ban she.


———. 2015. *CiNii books: Metadata and API: CiNii books OpenSearch for books and journals.*


doi:10.1080/01639374.2012.681605


In Search of Hong Kong Literature in a Digital Hub:
An Evaluation of Hong Kong Literature Database
from the Perspective of Researchers

Chui Yu Cheung, Dickson K.W. Chiu
Faculty of Education, The University of Hong Kong, Hong Kong

King Kwan Li
The Chinese University of Hong Kong Library, Hong Kong

Introduction

The Hong Kong Literature Database (HKLD) was launched by the Chinese University of Hong Kong (CUHK) library in 2000. It is the first digital database on Hong Kong literature, which serves as a research information tool for users to look for Hong Kong literature in the digital age. It is one of the early locally developed digitization projects in the field of Hong Kong studies. In particular, for the HKLD definition of "Hong Kong literature," CUHK library takes the most liberal definition that the literature is published in Hong Kong or written by authors living in Hong Kong. As the librarians are aware of identity politics well and note the evolving definition of "Hong Kong people," HKLD would like to show a holistic picture of Hong Kong literature and position HKLD as a neutral platform.

The CUHK Library has spent enormous resources on the establishment, improvement, and maintenance of HKLD to fulfill different user information needs. The CUHK Library conducted a quantitative user survey in 2007 to examine the profiles, purposes, and browsing habits of HKLD users (Hong Kong Literature News, 2007). The survey showed that users visited HKLD for learning (37%), research (30%), personal interests (27%), and teaching (5%). However, there has been no further follow-up study of HKLD by neither the CUHK library nor other scholars over the past decade. In recent years, more Hong Kong scholars and writers have been promoting Hong Kong literature and striving to set up a permanent Hong Kong Literature Museum (The House of Hong Kong Literature, 2018). Therefore, HKLD should be further enhanced in order to deal with increasing potential users and diversifying needs. Thus, it is worthwhile to re-examine HKLD from the perspective of both users and librarians.

Based on eight interviews with researchers of the CUHK, this article examines if HKLD can continue to meet the needs of researchers of Hong Kong literature studies successfully. The paper finds that first, HKLD serves as a hub of resources of Hong Kong literature to preserve Hong Kong literature to overcome the data collection problem for researchers, facilitating research with its comprehensive collection. Second, HKLD is a one-stop digital platform that indexed and digitized literary items for instant retrieval so that researchers can collect data in a more effective and efficient manner.
The new field of Hong Kong literature

Owing to low awareness of preservation and weak institutional support, information of Hong Kong literature had not been preserved very well. This was due to the ignorance of local culture in the colonial period, late formulation of Hong Kong identity, and the young history of Hong Kong literature as an academic discipline (Lo, 1988). Little attention was paid to Hong Kong literature in civil society and academy before the 1980s; thus artifacts of Hong Kong literature, especially those published in the early 20th century, were scattered in different places or even lost. Libraries have contributed efforts to the building of special collections on Hong Kong literature in recent 20 years only. Hong Kong public library set up Hong Kong literature collection in 1994, Fung Ping Shan library of the University of Hong Kong started to collect items on Hong Kong literature in the 1990s, while the CUHK library established Hong Kong literature collection in 2000. (Lo, 1988, 2005).

Moreover, there was a debate on Hong Kong literature as a field in the 1970s or before, as some scholars, from Taiwan and Mainland China, doubted if there was any serious literature in Hong Kong (Lu, 1988). It takes time for scholars to establish the legitimacy of Hong Kong literature as an academic discipline. There is a vibrant discussion of the definition of Hong Kong literature in the 1980s and onwards (See Wong, Lo, Tay, 1998; Cheung and Chu, 2002). The boundary of Hong Kong literature is vague due to the intimate interactions and high mobility of writers of Hong Kong, China, and Taiwan. Many writers traveled from mainland China, stayed in Hong Kong for some years, and then returned to mainland China or went to other places (Chen et al., 2018). Hong Kong was a freed port as compared to the political instability in mainland China and Taiwan, and anyone can publish their work in Hong Kong. Therefore, it is difficult to define clearly whether these writers are Hong Kong writers or not. Criteria of Hong Kong literature may depend on the birthplace of writers, place of publications, subject matter of literature works, and “consciousness” of Hong Kong identity. Such ambiguousness of Hong Kong literature has caused difficulties to confine the scope of Hong Kong literature, and made it a difficult task to index or collect Hong Kong literature and eventually institutionalizes as an “official” version. Further, the young discipline of Hong Kong literature is growing rapidly and it takes years to consolidate the field in the academy. It also takes time to develop supporting tools for research on this subject.

As a result, in the early days of Hong Kong literature studies, researchers encountered great difficulties in tracing materials of Hong Kong literature. Researchers realized their information needs of Hong Kong literature, such as works by a writer or works published in a particular period; however, but there was no index or bibliography of Hong Kong literature. There was only scattered collections of literary writers published in Hong Kong due to limited resources and interests. Many literary writings were published in different channels of newspapers, supplements, magazines, and journals. Further obstacles in searching exist as many Hong Kong writers made a living by writing popular fiction for different newspapers under different pen names, and submitted their “more serious” writings to journals with other pen names.
Worse still, researchers may not be able to get access to publications from various libraries, as they only started to build special collections on this subject since the 1990s. Many researchers often need to build their private collections of Hong Kong literature with their own resources, which means these materials could not be accessible to other users.

**The Hong Kong Literature Database**

Since the 1980s, there has been an emerging interest in Hong Kong literature in local universities (Wong, 1987; Lo, 1988; Cheung et al., 2002). Undergraduate courses on Hong Kong literature have been offered widely in universities in Hong Kong since 1999, and the Hong Kong Literature Research Centre was established at the CUHK in 2001. Hong Kong literature studies were institutionalized in Hong Kong in the 2000s, and this is an emerging academic field comparing to traditional Chinese literature studies and history.

In response to the growing information needs of Hong Kong literature, the CUHK library set up HKLD in 2000 with three main objectives: “(1) to serve the teaching and research needs of the university community and beyond on Hong Kong literature; (2) to provide easy access to materials on Hong Kong literature anytime anywhere in the world; and (3) to promote Hong Kong literature as a subject discipline to a wider audience globally.” (Ma, 2005)

As of November 2018, there are over half a million items indexed in HKLD, including Hong Kong literary journal articles (318,516 items), book jackets and table of contents of monographs of Hong Kong literature (19,082 items), Hong Kong newspaper literary supplement articles (178,368 items), other newspaper articles (6,741), as well as theses and dissertations on Hong Kong literature (377 items). Further, the full text of 197,330 digitized items is available in HKLD. The primary collection of HKLD was donated by Prof. Wei-luen Lo, a prominent writer and scholar who had collected artifacts of modern Chinese and Hong Kong literature with personal resources and networks for more than 30 years. Then, the collection of HKLD was developed by CUHK library with further purchase of literature publications, copyright permission granted by publishers, and academic support by the Department of Chinese Language and Literature, CUHK.

Although the CUHK library has been collecting, preserving, and organizing information of Hong Kong literature for a long time, it was a new attempt for the CUHK library to set up a digital database with the subject of Hong Kong literature. In 2016, there were approximately 4,500,000 visits to HKLD on average with users from Hong Kong, mainland China, Taiwan, the United States, Europe, and other places (Hong Kong Economic Journal, 2016; Wen Wei Po, 2016). HKLD, as a subject database, is popular among users, as there are only “more than 200,000” visits to the website of Asian Art Archive with an online catalog of contemporary art documents in the region. (Asian Art Archive, n.d.). It is well-recognized that HKLD has provided an important information service to users who are interested in and conduct research on Hong Kong literature (Hong Kong Economic Journal, 2016; Wen Wei Po, 2016).

There is only a limited discussion on HKLD in the field of librarianship and information management studies. For example, Ma, Wong, and Lau (2005), staff of the CUHK Library, introduced the process of setting up HKLD as well as its design and metadata structure. Ma
and Chan (2008, 2010) explained how the Hong Kong Newspaper Literary Supplements Digitization Project contributed to the growth of HKLD. However, current literature on HKLD focuses mainly on preparation for digitization and the process of making the digitized material accessible to users, while the user experience of HKLD is yet to be explored.

**Digitization and academic libraries**

As a digitization project on Hong Kong literature of the CUHK library, HKLD responds to the growing research interests in Hong Kong literature and the bloom of digital libraries worldwide in this century. In this Internet era, users expect to retrieve information in a timely manner and get access to library materials without geographical and time limitations. Thus, libraries also go digital and provide information services to users via online platforms. A digital library refers to a managed collection of information in digital formats to provide information services accessible over a network (Chowdhury and Chowdhury, 2003; Arms, 2000). Many libraries have recently become hybrid libraries to serve digital and printed information resources in an integrated information service accessible locally as well as remotely (HyLife, 2002). The following discussion focuses on the practice of digitization only and other features of digital libraries will not be covered.

With advanced information technology, libraries, museums, and other information organizations are able to preserve cultural and intellectual resources and provide online access of these resources to users via digitization (Deegan and Tanner, 2004; Lo, Chiu, and Cho, 2017; Lo et al., 2019). Hughes (2004) defines digitization as a process by which analog contents are converted into a binary code of 0s and 1s to be computer-readable. She invited readers to evaluate the costs and benefits of digitization with reference to information organizations’ needs and introduced stages of implementation of digitization projects. Though she recognized both time and financial costs of digitization, she also emphasized the advantage of digitization projects, including providing broader and enhanced access to a wider community, supporting the preservation, supporting collection development with collaborative digitization initiatives, and benefits for research and education (Hughes, 2004). On the other hand, Calhoun (2014) described digitization as a process of converting a physical item into a digital representation so that digital contents could be easily compressed and possibly manipulated with computer programs. Much other discussions on digitization have been made mainly from the perspective of information managers, especially on planning and implementation (Lee, 2001; Hugh, 2004; Rikowski, 2010).

On the other hand, it is essential to examine the user experience of digital databases with digitized items. After all, digitization projects are to preserve cultural and historical information for users of the present and future times (Lo, et al., 2017). Users play a prominent role in digitization projects, as they can give valuable feedback for the improvement of digital libraries for information organizations. It is always important to learn the information need of users so that information organizations can review their collection policy properly. Further, information-seeking behaviors of users can provide hints to enhance the design and features of digital databases.
Last but not least, it is equally important to identify the needs of specific groups of users. Carr (2007) pointed out that users do not consist of a conveniently homogeneous group, and there may be contradictory requests from different users, such as undergraduates, taught postgraduates, research students, full-time scholarly researchers, and external users from diverse professions. In particular, Brophy (2005) identified research students and research staff as highly demanding information users. As their research has to be original and presumably cutting edge, they usually request for in-depth coverage of a highly specialist area and timely information. As such, in this article, we focus on the researchers’ needs of Hong Kong literature studies for HKLD.

Objective and methodology

In this project, we conducted qualitative research on the user experience of HKLD through eight semi-structured interviews with researchers from two academic departments at the CUHK, who have been searching for information for research from HKLD. Semi-structure interview refers to a number of predefined questions are always asked in the interview, but subsequent questioning will depend on the response given (Bawden and Robinson, 2012; Pickard, 2007). With such flexibility, the research could explore the information needs of Hong Kong literature studies and user experience of HKLD covering depth and breadth.

As shown in Table 1, the background information of interviewees, including age, position at the University, year of usage of HKLD and research area, was presented for a general understanding of researchers. Six research students and two full-time teaching staff from the Department of Chinese Language and Literature, and the Department of Cultural and Religious Studies, CUHK were interviewed. Besides four researchers on Hong Kong literature, four researchers in other humanity fields are interviewed to cover more diversified needs, including Chinese Modern literature, Hong Kong culture, and Hong Kong cinema.
Table 1. Interviewee demographics.

<table>
<thead>
<tr>
<th>Interviewee</th>
<th>Status</th>
<th>Year of using HKLD</th>
<th>Research area</th>
<th>Information needs</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>M. Phil. student</td>
<td>6</td>
<td>Hong Kong Literature</td>
<td>Literary writing by a writer; Review and criticism on a writer</td>
</tr>
<tr>
<td>B</td>
<td>Ph.D. student</td>
<td>10</td>
<td>Hong Kong Literature</td>
<td>Literary writing by a writer; Review and criticism on a writer</td>
</tr>
<tr>
<td>C</td>
<td>Ph.D. student</td>
<td>1</td>
<td>Hong Kong Culture</td>
<td>Literary writing with descriptions of a social phenomenon in Hong Kong in the 1960s-1970s</td>
</tr>
<tr>
<td>D</td>
<td>M. Phil. student</td>
<td>4</td>
<td>Hong Kong Cinema</td>
<td>Review and criticism of Chinese cinema in literary and cultural magazines in Hong Kong</td>
</tr>
<tr>
<td>E</td>
<td>Ph.D. student</td>
<td>6</td>
<td>Hong Kong Literature</td>
<td>Literary writing by a writer; Review and criticism on a writer</td>
</tr>
<tr>
<td>F</td>
<td>Teaching staff</td>
<td>6</td>
<td>Hong Kong Culture</td>
<td>Records, reviews and criticism of Xiqu performance in literary and cultural magazines in Hong Kong</td>
</tr>
<tr>
<td>G</td>
<td>M. Phil. student</td>
<td>4</td>
<td>Hong Kong Literature</td>
<td>Literary writing by a writer; Review and criticism on a writer</td>
</tr>
<tr>
<td>H</td>
<td>Teaching staff</td>
<td>16</td>
<td>Chinese Modern Literature</td>
<td>Literary writing by a writer; Review and criticism on a writer</td>
</tr>
</tbody>
</table>

It is admitted that the sampling of qualitative research is limited and thus this research may not be able to cover the usage of HKLD by “general” users. However, this qualitative research could still reveal invaluable details of user information needs and information-seeking of HKLD. Further, the selected interviewees comprised junior researchers and senior researchers, ranging from M. Phil. students to full-time teachers. Thus, they had different years of experience of using HKLD, ranging from 1 to over 10 years. As some interviewees had been using HKLD since undergraduate studies, they could also comment on the evolution of their own needs as well as HKLD.

Results

HKLD as a resource hub for Hong Kong literature

As discussed in the previous session, information of Hong Kong literature had not been preserved very well due to low awareness of preservation and weak institutional support during the colonial period. Furthermore, the ambiguousness of Hong Kong literature has caused difficulties to confine the scope of Hong Kong literature, and made it difficult to index or collect Hong Kong literature and eventually institutionalizes as an “official” version.

All the interviewees appreciated HKLD as a hub of resources of Hong Kong literature and they could fulfill their information needs with the usage of HKLD. As the interviewees understood that many literary items they needed were indexed or accessible in HKLD, they
would “naturally” look for information of Hong Kong literature when they needed something from the Hong Kong literature.

Interviewees expressed that a lack of information tools and hub of resources of Hong Kong literature had traditionally been an obstacle to academic research in this field. Without HKLD, it would be difficult for researchers to develop research projects if they were not sure whether their required information of Hong Kong literature was available at all. They further commented that data collection and retrieving information of Hong Kong literature was more difficult than their counterparts in the field of Chinese modern literature, which had well-developed information tools for a long time.

As such, interviewees consider HKLD as one of the key information sources for their research project of literature and related studies because of its comprehensive collection. They appreciated that full text of more than 60 journals is available in HKLD, including *The Chinese Student Weekly* (中國學生周報) and *Hong Kong Literature Monthly* (香港文學). The interviewees were grateful that they could enjoy the comprehensive collection of HKLD. They found it satisfactory to search for Hong Kong literature in this digital hub and learning the historical and current development of the field.

As journals of Hong Kong literature were not well-preserved in public libraries and academic libraries in the 1990s and before, researchers might not be able to get access to these journals for research purposes without HKLD. With HKLD established in 2000, journals of Hong Kong literature and other literary items are available to users for a better understanding of the literary and cultural scene of Hong Kong in the 20th century. Thus, HKLD, as a hub of Hong Kong literature, lays a foundation of Hong Kong literature studies for researchers and interested members of the general public.

Mr. Ma, CUHK librarian and a member of the working team of HKLD, expressed that HKLD has taken an inclusive definition of Hong Kong literature, in order to preserve as much material of Hong Kong literature as possible for users of the present and future. The collection policy of HKLD is based on the place of publication of items of Hong Kong literature; in other words, HKLD provides a comprehensive collection of literary items published in Hong Kong, regardless of nationality, birthplace or residency of writers (Ma, 2016). This collection policy is a clear and concise one.

Further, the working team of HKLD comprises CUHK librarians and academics of Hong Kong Literature Research Centre, CUHK. Thus, the selection of information for HKLD is advised by experts in Hong Kong literature, while the information organization is supported by professional CUHK librarians. More importantly, Prof. Wei-luen Lo, a well-respected pioneer of Hong Kong literature studies, has collected items of Hong Kong literature since the 1970s. Prof. Lo donated her own collection and index of Hong Kong literature, which served as a foundation of HKLD. Hence, the collection of HKLD is considered to be comprehensive with academic standards by interviewees. Interviewee A commented on the collection, “We all know Prof. Wei-luen Lo donated her collection and took part in setting up HKLD. With her expertise in (Hong Kong) literature studies, we can count on her to select Hong Kong
literature materials. Of course, HKLD cannot cover everything, but Prof. Lo must have identified key items for HKLD.”

Meeting user information needs in a one-stop digital platform

As the items from different information sources are digitized and made accessible in HKLD, all the interviewees appreciated the convenience to get literary items published in newspaper supplements, journals, and monographs in a single platform, so that they needed not spend time on searching for information in multiple platforms or catalogs. This enhanced their efficiency in data collection. We further classified their information needs as listed in Table 2, which comprised three main types: (1) literary writing and writers, (2) cultural activities, and (3) literary description of a social phenomenon in Hong Kong. Next, we analyze how the functions of HKLD satisfy user information needs.

Table 2. Information needs of interviewees related to HKLD.

<table>
<thead>
<tr>
<th>Theme</th>
<th>Coding</th>
</tr>
</thead>
</table>
| Information related to writer(s) and his/her work | -when and where a writer publishes his/her literary writing  
                                      -interview and introduction of the writer  
                                      -review and criticism on literary writing  
                                      -source of literary writing and criticism, such as the title of literary magazine and year of publication |
| Information related to cultural activities | -records of cultural activities in newspaper and magazines, such as Chinese films and Xiqu performance  
                                      -review of cultural activities |
| Information related to a social phenomenon | -literary description of a social phenomenon in Hong Kong in the 1960s-1970s |

HKLD search functions

Users could retrieve information with the search engine provided by HKLD via “keyword search” and “search by author.” Keyword search is very straightforward. Basically, users could enter a keyword and then a list of items with the supplied keyword is shown. The article titles, author names, publication title, and publication date have been indexed in HKLD. Users could retrieve information at the article level in HKLD. For example, with the keyword of "Xi Xi" (西西), a renowned Hong Kong female writer, HKLD retrieved 327 items, published from 1965 to 2018, listed in the search result (see Figures 1 and 2). This is a very user-friendly function for users who are new to the field of Hong Kong literature.

“Search by author” provides more advanced functions. Users could search by author name, including real name and writer’s name. HKLD has identified authors of Hong Kong literature with their real names and pen names in Chinese and English, and literary items are also classified by type of writings and sources, including literacy text, review, newspaper clipping, monographs, journal article, and academic thesis. In the case of "Xi Xi", users could locate the author’s writing published in newspapers (16 items), journals (371 items), and monographs (44 items), and also identify literacy review on Xi Xi’s writing (251 items) (see Figure 3).
Figure 7. Keyword search interface.

Figure 8. Keyword search results.
Advantages of search function in HKLD

Obviously, there are advantages that HKLD allows users to retrieve different types of literary items in a one-stop digital platform. First, it is more efficient to search for information on a single platform. Second, users, with or without literature studies training, are able to locate literary items with the search engine provided by HKLD.

For those researchers who have literature studies training and look for literary information from specific information sources, they get access to the desired literary items in HKLD more quickly. For example, if a researcher would like to find a specific article published in a journal in a particular year, the researcher could browse all the issues of journals published in this particular year for the item without the need to access the physical copies. HKLD further supports browsing journals and newspapers in paper or microfilm format. However, the interviewees expressed that they did not prefer reading items in microfilm format as this would be very time-consuming.

For those who are new to the field of Hong Kong literature, they can retrieve literary items from the search engine of HKLD easily. As metadata of literary items are indexed in HKLD, users can simply type a keyword and then start searching even when they have no idea of the journal title or author who might have written about a phenomenon or a concept.

Interviewees expressed their difficulties in locating literary items before using HKLD, as they did not know how to start identifying information sources of literary writings. Interviewees pointed out that as Hong Kong literature is a young discipline, collections of Hong Kong writers were not complete. Complete collection of a writer would be more useful for researchers to understand the writing journey of a writer by tracing their work. As a result, researchers had to learn tacit knowledge of Hong Kong literature by themselves. It would take a long time for them to learn key sources of literary publication, such as literary
and cultural magazines in different periods and network of literary writers. An obstacle in locating literary information was that the time span of literary magazines was often short, such as two years or so, owing to the limitation of publishing resources. In the past, researchers could only learn such tacit knowledge slowly, either from their teachers and colleagues in classes or seminars, or from cited literary information in academic articles. Only through that could they know the required title of literary magazines or key discussions of the subject for retrieval.

With HKLD, researchers could make use of the search engine to retrieve records of literary items and follow previous discussions of a particular writer more effectively. HKDI provides online access to digitized items, or information to direct users to access them physically at the CUHK Library. Further, users could retrieve the metadata of related items for citation, including article title, author, and source (title, date of publication and page number). As a result, interviewees could fulfill their information needs for their research with usage of HKLD, directly and indirectly.

In general, all interviewees expressed their appreciation of HKLD, which is the first and largest digital database on Hong Kong literature. They are especially grateful that comprehensive and well-selected collection with many “old materials” published in the 1970s or before are available, and full text of more than 60 digitized journals are accessible in this digital hub. They also acknowledge that the search function of this one-stop platform is very helpful to their research projects. HKLD makes the “unwritten knowledge” of Hong Kong literature open to all users by providing a searchable index of literary items.

Recommendations for HKLD

Table 3 summarizes some suggestions from the interviewees for the further improvement of HKLD: (1) enrichment of collection, (2) upgraded features, (3) system stability, and (4) promotion and user education.

<table>
<thead>
<tr>
<th>Enrichment of collection</th>
<th>Providing up-to-date literary items (2014 or onwards)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Collaboration with other universities for filling the gap of collection, especially items published before 1949</td>
</tr>
<tr>
<td>Feature</td>
<td>Enhancement of search engine (with Boolean search)</td>
</tr>
<tr>
<td></td>
<td>Strengthening linkage between the online catalogue of the CUHK library and HKLD</td>
</tr>
<tr>
<td>System</td>
<td>Ensuring stability of user access to HKLD</td>
</tr>
<tr>
<td>User education</td>
<td>More promotion of HKLD in library orientation sessions.</td>
</tr>
</tbody>
</table>

First of all, interviewees recommended the CUHK library to enrich the collection of HKLD. They all noted that Hong Kong literature is an expanding field and thus they hoped that the CUHK library could put more resources on collection updates, both newly published materials and recently re-discovery of old materials. They suggested the CUHK library to provide more updated literary items in HKLD, especially items published on or after 2014.
They also encouraged the CUHK library to collaborate with other universities for filling the collection gaps of HKLD, especially items published before 1949. They also recommended HKLD to provide more literary items with full text online. Then, interviewees also recommended the CUHK library to improve the functions of HKLD. They suggested upgrading the search engine of HKLD with Boolean search and a stronger linkage with the CUHK library catalog.

As mentioned, Prof. Wei-luen Lo was responsible for selecting materials for HKLD at the beginning, and she is no doubt an expert of the field, but now we can see the scope of Hong Kong literature is expanding, thus HKLD has to take care the diversified information needs of more and more potential users. Since digitization and database setup is a time-consuming and costly process, without user participation may result in project failures (Warwick, et al., 2012). Therefore it is suggested that the CUHK library should include more users’ participation regarding the scope of the collection and the upgrade of features in different stages of the project. CUHK library may use inquiry records, surveys, and software such as Google Analytics to study the users’ information behavior (Bantin and Agne, 2010). Some other suggestions for user need studies include using deep log analysis, case studies, and longitudinal studies such as discussion in workshop, etc. (Warwick, 2012)

Further, interviewees also suggested the CUHK library to improve the system stability of remote access to HKLD, as they expected to get online access to HKLD anytime and anywhere. The stability of the system of HKLD is nowadays of great importance.

Lastly, it is recommended the CUHK library should do more promotion and user education to research postgraduates and new teaching staff. Although CUHK library has created a LibGuide of Hong Kong Literature (https://libguides.lib.cuhk.edu.hk/HKLit) to highlight HKLD and other relevant research resources, and has set up “CUHK Golden Jubilee Celestial Civilian Scholarship on Hong Kong Literature” to encourage CUHK students to use the HKLD collection for research, from the interviews it seems quite a lot of teachers and students had no idea about these efforts. It was inadequate for researchers on Hong Kong studies to learn just from classmates and colleagues about HKLD providing access to key literary and cultural magazines in Hong Kong. They suggested that more promotion of HKLD should be done so that more researchers, especially those in the field of humanities, could make better use of HKLD.

Conclusion

This is the first qualitative study to evaluate HKLD from the perspectives researchers in the humanities seeking for information on Hong Kong literature and culture from HKLD. This study found that HKLD could fulfill most of the information needs of the researchers. The interviewees generally appreciated the importance of HKLD as a unique one-stop platform providing multiple information sources, with a comprehensive collection, a user-friendly search engine, and a high academic standard. However, they also expected the CUHK library to make further improvement of HKLD in areas of collection, features of the search engine,
and system. They also recommended the library to do more promotion and user education on HKLD.

HKLD provides online access to Hong Kong literary items to local and overseas users for free. With HKLD, the literary items collected by Prof. Wei-luen Lo and the CUHK library could be circulated among a much larger community. It does provide significant support to researchers in search of Hong Kong literature in this digital hub.

As for continuing research plans, we are interested in further qualitative research on the usage of HKLD by other user groups for a better understanding of the role of HKLD in teaching, learning, and research in related fields. We are also considering plans to utilize the extensive data in the HKLD for digital research. For example, we know literature researchers have long been studying authors’ writing styles and social networks. Text mining and analysis, spatial analysis with GIS, and data visualization, etc., might become powerful tools to help them expand their research scope and enhance their research quality of Hong Kong literature.

Acknowledgment
We thank Dr. Patrick Lo for his advice on this study and suggestions for our future research directions.

References


Asian Art Archive, No Date, Retrieval from http://www.aaa.org.hk/About/Overview.


"Life achievement award: Lo Wai-luen preserves local literature and culture" (終身成就獎──盧瑋鑾(小思)傳承本地文學 in Chinese), *Hong Kong Economic Journal*, 2016-04-26, D01.

"Speech at award ceremony: Lo Wai-luen determines to keep records of Hong Kong History" (領獎不忘諫言 小思誓史存照 in Chinese), *Wen Wei Po*, 2016-04-30, B07.

Reported by Charlene Chou, CTP Chair

Since the article entitled: “Emerging Technical Services: The Vision of the Committee on Technical Processing” was published in the Journal of East Asian Libraries, no. 167 (2018), the Committee on Technical Processing (CTP) Chair has received positive feedback from several CEAL members. This report is Part Two and focuses on the progress of CTP activities and work plans to fulfill the goals of CTP in the past year.

Demand-driven training

After the 2018 cataloging workshop, CTP received requests to provide more training on emerging topics such as cataloging special materials/unique resources, manuscript and archives, linked data, BIBFRAME, identity management/authority control, e-resources metadata, proposing new LC subject and genre terms, non-MARC metadata, managing digital humanities/digital scholarship metadata and controlled vocabularies, etc. In the 2019 CEAL Cataloging Workshop, we included some of these topics. For example, “3R (RDA Toolkit Restructure and Redesign) Update” was presented by Kathy Glennan, the RDA Steering Committee Chair, since the beta is available for testing and the Toolkit has been transformed from an e-book for RDA rules to a data dictionary for both cataloging/metadata librarians and developers. Another highly demanded topic, “Cataloging special materials,” was taught with hands-on exercises and cataloging examples of artist's books, 3D objects, visual materials, e.g. posters, postcards and so on. The participants were actively engaged in small group discussions and exercises.

In light of linked data, authorities, faceted vocabularies and identity management have been crucial for library metadata management. As usual, we had reports and updates from the CJK NACO, BIBCO, CONSER and SACO funnel projects, as well as ERMB cooperative cataloging for e-resources project. Since NACO CJK Funnel References Project for authority records was just launched, workflows and policies were explained thoroughly.

What does data curation mean to library technical services? This presentation was given by a data scientist, Matthew Mayernik from NCAR (National Center for Atmospheric Research Library), and key topics included data curation initiatives within libraries, data within library collections and distributing library metadata beyond libraries. The presenter emphasized the importance of authorities/identity management done by cataloging community, and recommended enhancing more authority work/controlled vocabularies for data curation. NCAR implemented a project to launch its DASH (Digital Asset Services Hub) Search, which allows user to find and access digital assets created and published by NCAR community programs, and this Hub distributes library metadata beyond libraries. Since
Google just launched its new Google Dataset Search, strategies/methods to make more library metadata discoverable through this Google search engine would be an important next step.

Enhancing the discovery of unique and digital collections through collaboration between public and technical services

In addition to providing quality metadata for e-resources, unique objects and digitized collections, there are increasingly enormous demands for DH/DS (Digital humanities/digital scholarship) projects. The Technical Services Librarians are committed to provide quality metadata and to ensure that these resources are findable, shareable and sustainable in the networked environment worldwide. The 2019 CTP meeting explored and examined how to provide and maintain library metadata for better discovery of these resources, such as using controlled vocabularies in the context of the linked data environment. Most importantly, we need to know how to communicate and collaborate with diverse stakeholders to fulfill the goals and implement projects in terms of user-centered services. Three presentations demonstrated stories and testimonies for using a collaborative approach to improve the discovery of digital collections, unique and rare materials.

The video presentation of a new DH Platform—Academia Sinica Digital Humanities Research Platform—was played and a wrap-up discussion, “What Roles Will the Cataloging/Metadata Librarians Play for DH/DS Projects,” was conducted. We strongly believe that collaboration makes global digital libraries sustainable due to shared cost and expertise and that library portals should be connected to international frameworks, information hubs and regional centers such as IIIF and DARIAH, etc. Quality metadata are mandatory to enhance the discovery of diverse resources by using controlled vocabularies and international standards. Most importantly, we embrace emerging technology such as linked data and use authoritative and linkable thesauri such as VIAF, AAT, LCGFT, etc. rather than reinventing the wheel to create a silo project with unlinkable metadata.

DH/Digital scholarship resources for East Asian studies and beyond

Some CEAL colleagues attended the International Conference on the Cyberinfrastructure for Historical China Studies at Harvard China Center, Shanghai in March 2018, the Digital Humanities Asia 2018 Summit at Stanford in April and the Eighth Conference of Japanese Association for Digital Humanities (JADH2018) “Leveraging Open data” in September 2018. Based on the resources of these conferences and others, CTP has compiled a list of Digital scholarship resources for East Asian studies and beyond and posted it on the CTP website. For the direct link to the document, go to http://www.eastasianlib.org/ctp/Projects/2017-2020/Digital_Scholarship_Resources_CTP.docx.

After this bibliography was posted on eastlib listserv, we received several emails from the US, Europe and Asia to support sharing this bibliography in the CEAL community. We have received several comments to update links for new versions of tools/platforms. On October 4th, 2018, we received an enquiry forwarded from a scholar in Japan who would like to see if CTP is planning to update/maintain this bibliography. “If CTP is planning to make...”
efforts to update this list in the future, his institution will not need to create their own list.” This was a great suggestion and encouragement that this bibliography gets attention from outside North America.

**In service to the future**

The theme of the 2020 CTP program will be “enhancing the discovery of East Asian resources with linked data for cultural heritage and digital humanities.” Semantic technology and linked data enable large-scale Digital Humanities research, collaboration and aggregation as well as Cultural Heritage (CH) data, which is multilingual, semantically rich and highly interlinked. There are large national and international CH portals, such as Europeana, to large open data repositories, such as the Linked Open Data Cloud, and massive linked library data in the US, Europe, and Asia in recent years. More East Asian libraries have implemented IIIF (International Image Interoperability Framework) for their digital collections to share, view and curate high-resolution images internationally. The 2020 CTP meeting will explore and discuss how to enhance the discovery of East Asian resources with linked data for cultural heritage and digital humanities and in light of cultural diversity. For instance, RISE (a technical infrastructure aiming to bring digital resources and research tools together in one platform for DH research) and CHMAP (a set of WebGIS tools to display maps from different sources, including IIIF images) will be presented.
In Memoriam

Dianna Ye Xu

I’m deeply saddened to share with you that Ms. Dianna Ye Xu, a former CEAL member and librarian at University of Wisconsin-Madison passed away in Eugene, Oregon on June 18th, 2019, after a brave battle against cancer. She was 64 years old.

A native of Beijing, Ms. Xu graduated with a BA in Chinese literature from Beijing Teachers’ College in 1982 and an MA in literature from Dartmouth College in 1993. Before coming to the U.S., she taught traditional and contemporary Chinese literature for nine years at a college in Beijing. Prior to University of Wisconsin-Madison, Ms. Xu worked as the Chinese Studies Librarian at the East Asia Library at University of Washington from 2003 to 2006. At University of Wisconsin-Madison, she worked first as Chinese Studies Librarian, then as East Asian Bibliographer, and most recently, as East Asian Studies Librarian. Dianna retired early this year. In recognition of her distinctive contributions to her profession and the university, on June 10th, 2019, the University of Wisconsin-Madison conferred emerita status to her. In March 2019, the faculty of Center for East Asian Studies at University of Wisconsin-Madison also presented a plaque of recognition to Dianna, on which it says: “In deep appreciation for your outstanding service.”

Dianna was a passionate and dedicated librarian, an active member of CEAL and a wonderful colleague. She will be missed by many in the East Asian studies and library community. On behalf of many of her colleagues at CEAL, I convey our sincere condolences to Dianna’s family and loved ones.

Submitted by Zhijia Shen
Institutional News

*University of Washington East Asia Library receives Tateuchi Foundation grant*

With deep gratitude I am delighted to share the good news that the East Asia Library (EAL) at the University of Washington has received a generous grant of $189,600 from the Atsuhiko and Ina Goodwin Tateuchi Foundation in support of its current projects over three years starting September 2019 through the 2021 fiscal year. These projects include space assessment and design for renovation of EAL’s main location to meet the new demands of the changing academy; workshops to support the research of graduate students in China, Japan, Korea, and Taiwan studies; preservation binding and processing for EAL’s newly cataloged CJK special collections and UW Libraries’ special photograph collection of Japanese Americans; further developing exchange programs with libraries in Japan through visits of librarians and library professionals; and travel grants for Japanese library science scholars to conduct research at the University of Washington. Focusing on UW Libraries strategic plan, these projects will significantly enrich our students’ experience, accelerate scholarship and learning, and enable our library to grow as a learning organization.

We express our heartfelt appreciation to the Tateuchi Foundation for the unwavering support. Thanks to their generosity, in the past years our East Asia Library has been able to accomplish transformational projects, such as projects to get rid of its decades-old Japanese cataloging backlog, to inspire support from other grants and community donors to eliminate its historical Korean cataloging backlog, and to significantly reduce its historical Chinese cataloging backlog, among many other significant projects.

With the current grant from the Tateuchi Foundation, we enthusiastically look forward to working with colleagues in and outside the UW Libraries and partners across the Pacific to carry out these projects.

Submitted by Zhijia Shen
New Appointments

Mitsutaka Nakamura began work as Washington University Libraries’ Japanese Studies Librarian on July 29, 2019. Mitsutaka holds a bachelor’s degree in political science from Waseda University in Tokyo, Japan, and a MSLIS degree from the University of Wisconsin-Milwaukee. He comes to the University Libraries from the C.V. Starr East Asian Library at Columbia University, where he served as a bibliographic assistant, focusing on Japanese acquisitions. Prior to his tenure at Columbia, Mitsutaka worked in various positions in film and media archives as a translator and researcher. He can be reached at m.nakamura@wustl.edu.

Submitted by Joan Wang
Publication News

“East View Information Services is pleased to announce the release of the Late Qing and Republican-Era Chinese Newspapers, an Open Access collection. This is the first of many collections to be built in partnership with the Center for Research Libraries as part of the East View Global Press Archive’s CRL Alliance.”

Submitted by Carolyn Fennell


Submitted by Cindy Ho