Leveraging NISO Standards and Best Practices to Improve Discovery and Access of Digital Resources

Bie-Hwa Ma, Shi Deng and Susan Xue

Abstract:

To tackle issues related to metadata and improve discovery of and access to digital resources, National Information Standards Organization (NISO) has developed a series of standards and recommended best practices. The impact of these standards has been discussed in depth. However, no research has been done in the areas of how those standards and best practices can be used for digital resources in Chinese, and strategies for implementing those standards. This paper aims to identify common issues in discovery of and access to Chinese digital resources; suggest feasible solutions; discuss how to leverage NISO standards and recommended practices; and recommend ways to promote the standards and best practices to stakeholders.

To cite this article:

To submit your article to this journal:

Go to http://ojs.calaijol.org/index.php/ijol/about/submissions

Leveraging NISO Standards and Best Practices to Improve Discovery and Access of Digital Resources

Bie-Hwa Ma
University of California at San Diego Library, San Diego, California, USA

Shi Deng
University of California at San Diego Library, San Diego, California, USA

Susan Xue
University of California at Berkeley, Berkeley, California, USA

ABSTRACT

To tackle issues related to metadata and improve discovery of and access to digital resources, National Information Standards Organization (NISO) has developed a series of standards and recommended best practices. The impact of these standards has been discussed in depth. However, no research has been done in the areas of how those standards and best practices can be used for digital resources in Chinese, and strategies for implementing those standards. This paper aims to identify common issues in discovery of and access to Chinese digital resources; suggest feasible solutions; discuss how to leverage NISO standards and recommended practices; and recommend ways to promote the standards and best practices to stakeholders.

Keywords: NISO standards, metadata, discovery, access, Chinese digital resources, licensing, vendor education

BACKGROUND

On one hand, the rapid growth of digital[1] resources and the emergence of information technology have enabled users to discover and access digital resources anywhere and anytime they need it. On the other hand, the increasing number of stakeholders and their products and services in the supply chain and the continually evolving technology and discovery tools have made effective access to digital resources more challenging. For digital resources in Chinese, the language and cultural barriers and lack of standardization have brought new challenges and made organizing and discovering digital resources even more complicated. To address content organization and discovery issues, the National Information Standards Organization (NISO) has been working

---

[1] This paper uses "digital" and "electronic" interchangeably, as the two terms have not been clearly distinguished and both terms have been used in standards and best practices.

closely with representatives from publishers, content providers, libraries, link service providers, software service providers, and other stakeholders in the digital resource supply chain to develop, establish, maintain, and promote a series of standards and recommended practices for creating and managing digital content and metadata. All of this is done with the goal of improving user experience with digital resources.

Much research has been conducted regarding the relationship between NISO Standards and discovery. Pesch (2014) discussed how the Knowledge Base and Related Tools (KBART) could improve OpenURL linking through the definition of best practices for exchanging title-level metadata, and its potential in improving discovery services. Marshall & Reynolds (2013) studied how Presentation and Identification of E-Journals (PIE-J) would address problems related to e-journal presentation. Kasprowski (2012) discusses how the KBART working group investigates and identifies OpenURL metadata issues in the serials supply chain. Lagace *et al.* researched on four standards and best practices recently released or currently underway at the NISO: KBART, PIE-J, Open Discovery Initiative, and Open Access Metadata Indicators. The applicability of these standards and best practices for different stakeholders of the information community were discussed (Lagace, Kaplan, & Leffler, 2015). The role of NISO standards and best practices has been discussed in depth, however, no research has been conducted on how those standards and best practices can be used for digital resources in Chinese, and what strategies can be used for implementing those standards. In fact, standards in presenting digital resources in China are very different from those in the United States. The lack of unified standards in producing, processing, and presenting digital resources has seriously hindered the access and discovery of those resources, both domestically and internationally.

This paper intends to focus on strategies for improving discovery of Chinese digital resources by examining these issues: identifying common issues that hinder users from discovering and accessing Chinese digital resources, discussing how to leverage NISO standards and recommended practices to improve discovery and access to Chinese digital resources, recognizing stakeholders' efforts in creating good records, and recommending strategies for promoting the standards and best practices to stakeholders through education, training, workshops, and incorporation of these standards and best practices into official agreements.

COMMON ISSUES WITH DISCOVERY OF CHINESE DIGITAL RESOURCES AND FEASIBLE SOLUTIONS

There has been little discussion about issues with access and discovery of Chinese digital resources in Chinese literature. A thorough search of China Academic Journals database, Wanfang e-journals database and Airiti indicates that research in this area has just begun and in-depth studies have not yet emerged. One relevant article was a summary of the digital library conference "2013 Library New Technology Forum - Discover the Ultimate Beauty of the Library Services," focusing on discovery of digital resources and held in Shanghai in 2013. The author discussed the fact that the quality and quantity of metadata, standardization of metadata, and the mechanism of discovering and connecting to full-text resources are critical to a well-functioned discovery system (Sun, 2013). In 2014, a conference focusing on discovery system and related technology was held in the National Library of China; the relationship between discovery system and metadata was one of the topics of the conference (Xin hua yue du, 2014). The impact of high quality metadata on discovery of Chinese digital resources was also discussed in a 2011 article. The article mentioned that cooperation between discovery system providers and content providers would be the best way to obtain the best metadata, so as to improve discovery of content (Sa, Wen, & Qu, 2011). Another article discussed in general that lack of standards in China in developing digital resources has caused "information island," integration of digital resources and their management has been

difficult which has a negative impact on accessing and discovering information (Zhai, 2014). An article published in 2013 discussed issues with discovering digital resources due to the lack of a database or storage of metadata for digital resources, and the lack of management of metadata, and how this has caused low quality in search results (Ouyang, 2013). From a different angle, another article discussed the dissemination issues of theses and dissertations published in Taiwan. It suggested assigning DOIs to improve the discovery, persistent access, and interoperability with other systems which ultimately disseminate theses and dissertations to the world in a much more efficient way (Zhang, 2014).

In English literature, searching the Library Literature & Information Science database yields very few articles related to metadata standards and discovery of Chinese digital resources. Yurong Y. Atwill compared and discussed some technical issues of three major e-journal databases from China (Atwill, 2005). Search results can be very different for terms used, in Chinese and/or English, and some articles are hard to find; it depends on the metadata or level of metadata indexed for keywords and/or full-text searching. Although all three databases allow browsing of individual journals, some provide only selected journals and do not provide a cover-to-cover surrogate for the print (Atwill, 2005). When discussing e-book development in China, Anthony W. Ferguson mentioned that the lack of display standards forced library patrons to learn multiple software front ends and remember which collection does what (Ferguson & Ko, 2004).

At a practical level, librarians have done some pioneering work. At the CEAL 2012 conference, Bie-Hwa Ma presented and discussed issues she faced when processing Chinese digital resources, such as incorrect metadata, incomplete or misleading content presentation, and the URL linking problem. She pointed out that the source of the identified issues could be traced back to any parties involved in the supply chain. She suggested a feasible solution is for all parties to comply with established standards and practices, such as OpenURL, KBART, and PIE-J (Ma, 2012). At the CEAL 2014 workshop, Connie Lam discussed metadata problems affecting access and discovery of Chinese digital resources. Lam (2014) pointed out that aggregators and content vendors have made great efforts in improving the quality of metadata and are willing to help librarians and library users to have better access to resources. Susan Xue gave a report on a survey conducted in January and February 2014 on metadata standards and best practices for Chinese e-resources. The survey results showed that the top five standards vendors/publishers followed were MARC 21, other classification, ISSN/ISBN, PIE-J, and LCSH. The top five standards vendors/publishers were interested in following were DOI, MARC 21, ISSN/ISBN, KBART, and OpenURL. The survey results also showed that vendors/publishers and librarians perceived some difficulties in promoting metadata standards, including the fact that complying with standards will increase product cost, and vendors have no metadata expertise to provide certain metadata or follow the standards (Xue, 2014).

Users found that, compared with digital resources in English, metadata and discovery issues with resources in Chinese are more serious. Libraries today spend significant library collection funds on digital resources, and we need to communicate the *value* of these resources to the academic community, to show that the investment is worthwhile and that the resources are supporting academic teaching and research. It is critical that we address these issues strategically in order to ensure maximum access and discovery. One of the most effective strategies is to develop and implement standards and best practices among all parties in the e-resources supply chain.

In the United States, NISO has been taking great leadership in working together with content publishers, libraries, and software developers to identify, develop, maintain, and publish technical standards to manage information in our increasingly digital environment.

There are many metadata related standards and recommended practices developed by NISO to support access and discovery. The following standards and practices will be discussed in this paper:

- Presentation and Identification of E-Journals (PIE-J)
- Digital Object Identifier (DOI)
- OpenURL
- Knowledge Based And Related Tools Recommended Practice (KBART)
- Open Discovery Initiative (ODI)

We strongly encourage publishers, content providers, libraries and all stakeholders in China to comply with and adopt these standards.

PIE-J: SOLUTIONS TO CONTENT AND METADATA PRESENTATION

The success of e-resource discovery and access systems relies heavily on the quality of contents and their metadata, which is based on how the contents and metadata are presented online. If content and its metadata are incorrectly or inappropriately presented, other stakeholders downstream would not be able to provide a good service for that specific e-resource as the information supply chain has been broken or weakened from the beginning.

Most of Chinese commercial e-resources are digitized from print journals and books. Listed below are the most common presentation issues of Chinese e-resources that have been seriously affecting the discovery and access.

- Contents not digitized from cover to cover: University of California (UC) e-journal providers, with the exception of Dragon source, do not digitize advertisements and front matters (title page, cover, masthead, colophon, etc.), which actually provide significant information for resource identification and metadata creation in discovery systems. Some providers include thumbnails of the cover images, but the resolution is too poor to display the content.
- Missing content (characters, paragraphs, and pages): This can be due to bad scanning workmanship or the lack of contents from the libraries' originals (see Figures 1 and 2).
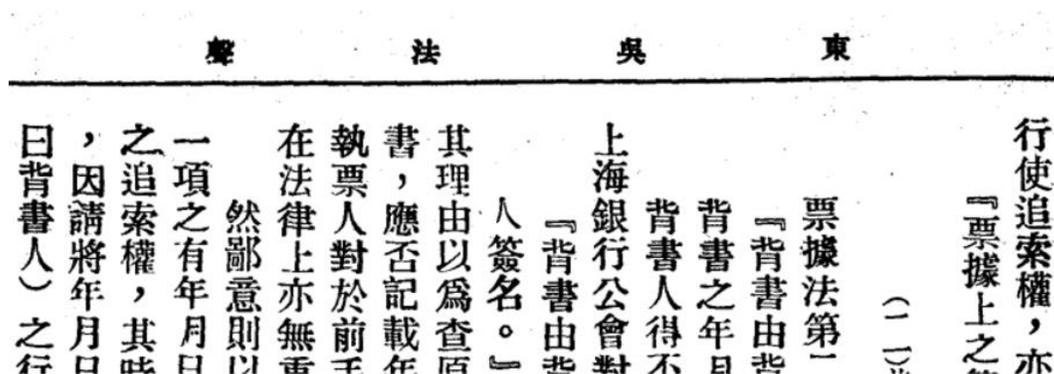

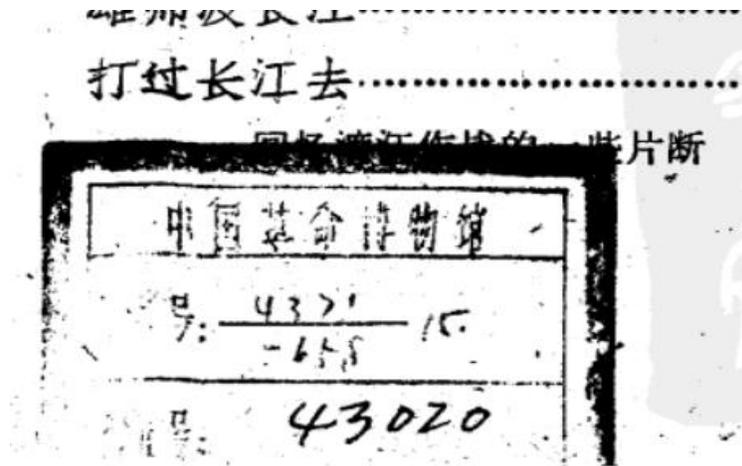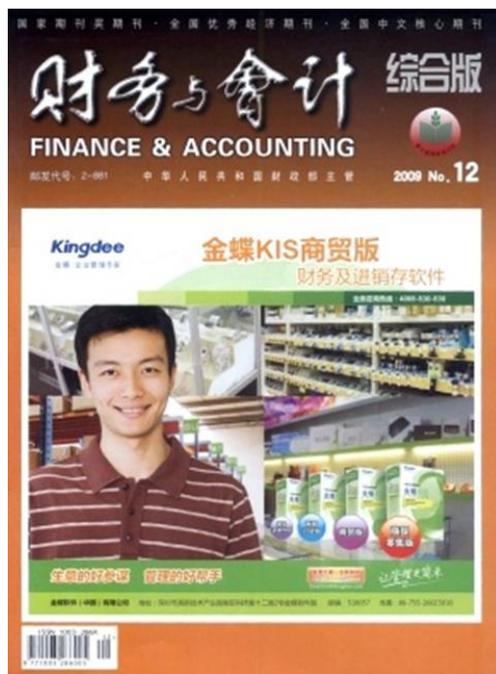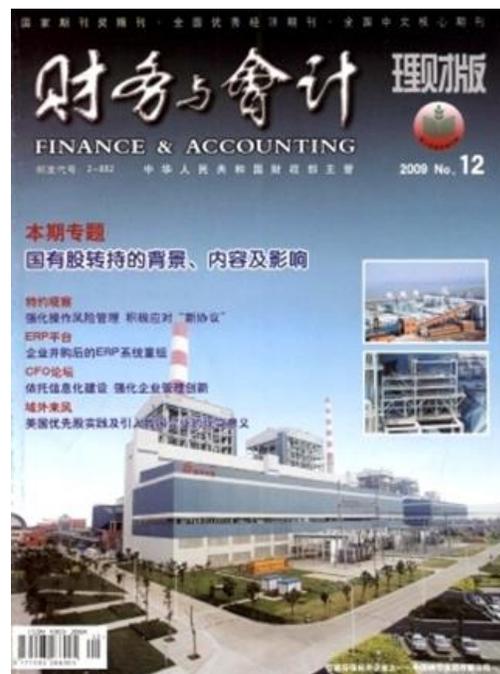Figure 1. Missing content resulted from bad workmanship

Figure 2. Content covered by the label of the holding library owning the original print material

- Renamed or renumbered content: Not only does this practice mislead users and affect link resolution and discovery in other forms (see Figure 3), but it also leads to incorrect citation (see Figure 4).

2009Year 23Issue   2009Year 24Issue

Figure 3. "Cai wu yu kuai ji" was published monthly in two parts, Zong he ban and Li cai ban, in 2009. The provider misnumbered the parts published in 2009, "no. 12" as "no. 23 & no. 24".

【Source】财务与会计 ,Finance & Accounting , Editorial E-mail ,2009(23)
【Source】财务与会计 ,Finance & Accounting , Editorial E-mail ,2009(24)

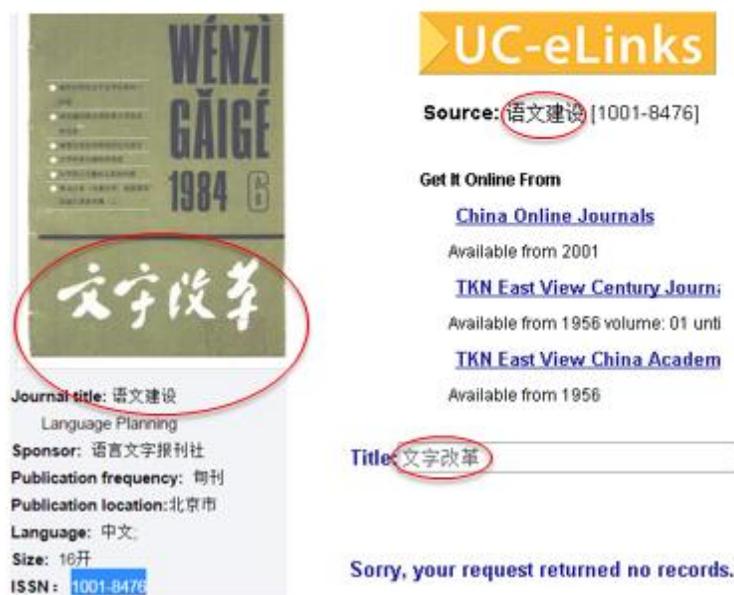| 刊名:	财务与会计（综合版）	PKU
Journal:	FINANCE AND ACCOUNTING	
年，卷(期):	2009, (12)	

Figure 4. Incorrect citation resulted from the renumbering of journal issues (top); the correct journal enumeration (numbering) metadata was provided by another package (bottom)

● Hidden title change practice: This is one of the biggest issues amongst UC-subscribed Chinese journal packages with the exception of Taiwan Electronic Periodical Services (TEPS/CEPS). Previous titles are displayed and represented under the latest title and the new ISSN; this means all the contents of previous titles are collated at the same website under the latest title and/or under the entry of the latest entry in the title list. As a result, previous titles are often not indexed in the knowledge base database, which has contributed to a large portion of OpenURL resolution failures and resulted in 'no hits' in the "A-Z journal title list" (see Figure 5). The biggest issue is that some providers incorrectly cite the articles of previous titles under the latest title (see Figure 6).

Title history of 拼音 [Pin yin]:

No ISSN	拼音	1956-1957
0511-4764	文字改革	1957-1985
1001-8476	语文建设	1986-

Figure 5. The former title "文字改革 [Wen zi gai ge]" was recorded under its later title "语文建设 [Yu wen jian she]" and a new ISSN on the 1984(6) issue landing page (left). This misrepresentation causes the former title to be indiscoverable in SFX KB (right) (viewed Mar. 21, 2016).
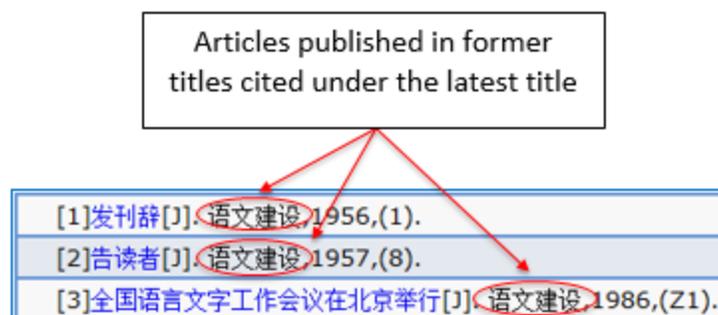
Figure 6. Although this longstanding citation issue was resolved by this provider at some point after March 2012, it remains a common issue among Chinese journal databases.

- Inaccurate ISSNs: The inaccurate ISSNs provided on the digitized resource or in the metadata are responsible for many OpenURL linking failures. Some Chinese publishers bought or borrowed an ISSN for their journal. This could be a result of the publishers having difficulties during the required procedure prior to the ISSN application. The publication would need to be pre-approved by General Administration of Press and Publication of the People's Republic of China, who would then grant the China domestic uniform serial publication number (CN), contingent for the application for an ISSN (ISSN China Center, n.d.).

To avoid discovery and access issues caused by incorrect ISSNs, vendors and publishers should verify ISSNs against the ISSN portal maintained by the ISSN International Centre or against the ISSN China Center portal which provides free searches by ISSN or journal title (ISSN Zhongguo guo jia zhong xin, n.d.). When providing metadata, record the ISSN that appears on the scanned copy of the original print material as "print ISSN," and record the one assigned for the online version as e-ISSN. However, do not record inaccurate ISSNs.

According to the new ISSN rule, "National Centres should assign an ISSN to both the digital reproductions and to the original print versions when the latter are not already identified" (International Standard Serial Number International Center, 2015, p. 18). This also applies to ceased print serials that are owned or not owned by publishers. The publisher who owns the content or the providing library/archive when the content is not owned by a current publisher can request the respective National Center to assign the ISSN(s). We are hoping that ISSN China Center will work with its governing agency for a feasible national practice to comply with this new ISSN rule which will enable serial contents to become more discoverable and accessible in the Internet era.

Presentation and Identification of E-Journals (PIE-J) Working Group was formed by NISO in July 2010. The group developed and published a recommended practice, *PIE-J: The presentation and identification of E-Journals* in 2013 (National Information Standards Organization [NISO], 2013). The recommendations have addressed the Chinese e-resources presentation and identification issues described above. Important highlights include:

- Retain content once it is published, whether digitized from the print or born digital. Do not remove, rename, or renumber content.

- Retain title and citation information under which articles were originally published

- Display:

    - title histories, including information relating to title changes and related metadata

     ○ correct ISSN for different formats and for changed titles

     ○ vital publication information across the history of a journal

   ● Create a presentation that allows easy access to all content

Complying with PIE-J recommendations would not only ensure the accuracy of important metadata, but also resolve the OpenURL failures and increase the usage of licensed resources.

PERSISTENT URL AND OPENURL: SOLUTIONS TO URL ISSUES

Linking technology has been a great asset since the Internet emerged in the mid 1990s by connecting users to e-resources. However, many studies, show that URLs in citations have been both growing and decaying rapidly (Gul, Mahajan, & Ali, 2014; Kumar & Manoi, 2012[2]; Prithviraj & Kumar, 2014). Gul et al. pointed out that the main causes of URL decay are the frequent substitutions of URLs and site death. Although some software tools have been used for identifying broken links since the turn of the century (Beam & Copeland, 2001), stakeholders including subscribing agencies and users would still spend hours co-troubleshooting, investigating, and/or updating broken linkage from different sites. In order to improve the URL stability and to reduce manpower maintaining URLs, new powerful technologies such as persistent URL and OpenURL were invented.

A persistent URL links users to a HTTP redirect server for retrieving the current actual URL and then seamlessly links to the target resource (see Figure 7). When a URL is changed, the persistent URL stays the same, while the URL address in the persistent URL registry is replaced with the new current URL. Examples include Online Computer Library Center's (OCLC) BibPURLs. Digital Object Identifier (DOI) link is also a persistent URL, but DOI systems[3] nowadays have more functionality (Corporation for National Research Initiatives, n.d.; International DOI Foundation, 2012).[4]

As defined by Crossref, the DOI registry center, a Digital Object Identifier (DOI) is "a unique alphanumeric string assigned to a digital object – in this case, an electronic journal article or a book chapter" (Crossref, n.d., "What a DOI is:" para. 1). A typical DOI link comprises the URL for the DOI resolver "http://dx.doi.org/" and a DOI. For example, http://dx.doi.org/10.6140/AP.9789865663933, which will be resolved to the current URL of the e-book, 中醫圖畫通說 [Zhong yi tu hua tong shuo] with the DOI "10.6140/AP.9789865663933." DOIs are assigned by DOI Registration Agencies (RA) which are authorized by the International DOI Foundation (IDF), the governing body of the DOI system. Unlike a regular URL, a DOI link does not change over time and no local link maintenance is needed because the URL of the DOI's corresponding object is maintained by its publisher or owner at the authorized registry. The DOI syntax became a national standard, ANSI/NISO Z39.84 (NISO, 2010c) in 2005 and an international standard, ISO 26324 in 2012 (ISO, 2012).

---

[2] From the study of 2890 URL citations cited in 689 research articles published in LIS open access journals from 1996 to 2009, Kumar et al. found that the percentage of URL citations increased from 5.98 percent in 1996 to 27.79 percent in 2009. However, 26.08 percent of all citations were not accessible during the time of testing.

[3] The DOI system's core is the 'Handle' system (Corporation for National Research Initiatives, n.d.), an open source software developed by Corporation for National Research Initiatives (CNRI).

[4] The updated DOI systems can have multiple resolution which allows one entity to be resolved to multiple pieces of data or entities. The systems can also work with Proxy servers and OpenURL system in which DOIs can be a parameter (element) embedded in an OpenURL, see chapter 3, Resolution of DOI handbook (International DOI Foundation, 2012).
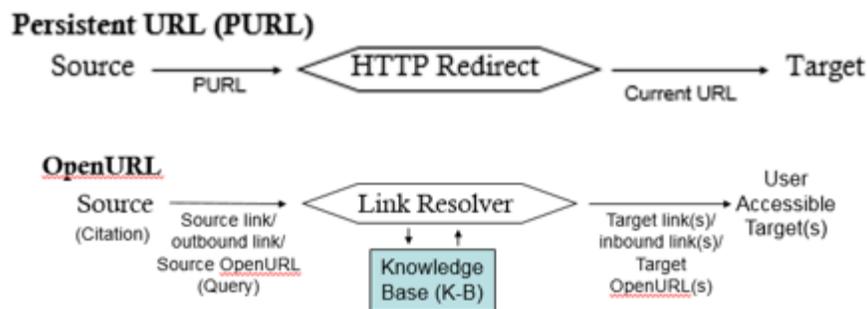
Figure 7. Persistent URL & OpenURL

OpenURL technology involves more components (see Figure 7).

1. It starts with a source link. It is created based on the metadata from the source/citation, the data about the user's affiliation, and the data about the location where the request was made. The source link is then passed to a link resolver for resolution.

2. By matching the metadata embedded in the source link against the metadata of resources, licenses, and restriction stored in the Knowledge Base (KB, a database created for linking OpenURLs to resources), the link resolver identifies the appropriate copies to which the user has access.

3. The target links are then created to link legitimate users to the resources.

The two links below are sample OpenURLs created by provider and UC campuses (UCs) respectively:

http://www.springerlink.com/openurl.asp?genre=book&isbn=978-1-58829-288-9

http://openurl.cdlib.org/?sid=SCP:SCP&genre=article&__char_set=utf8&issn=1089-8689
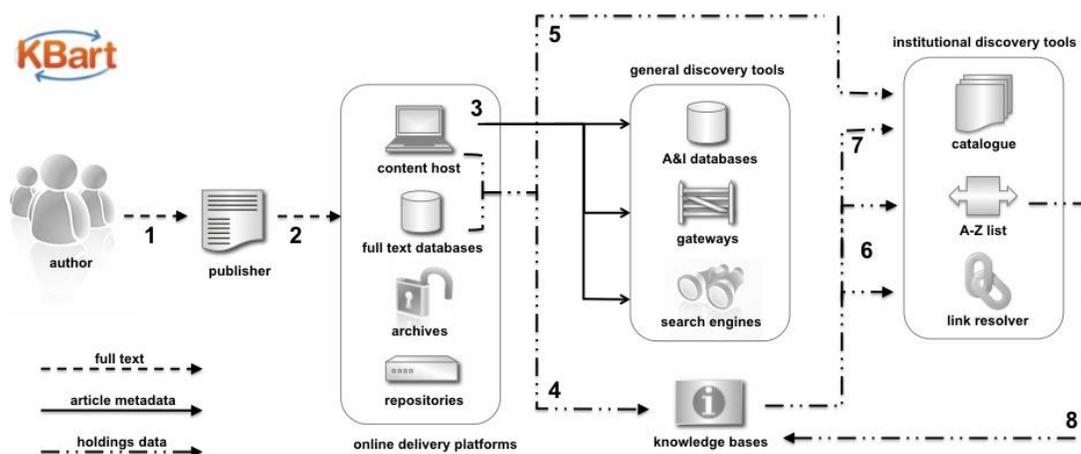
Implementing the OpenURL framework has many benefits such as providing "appropriate copies" of a referenced content, using dynamic and flexible links to reduce linking failures, lowering the cost for links maintenance, and increasing interoperability for data exchange, especially when the content is moved around between different providers and/or platforms (NISO, 2010b). In fact vendors do not necessarily have to implement the OpenURL system if their clients have subscribed to OpenURL link services, which most libraries do. Therefore, in most cases, what vendors need to do is to regularly provide accurate, comprehensive, rich, and current metadata in KBART template (see section 2.3) or in MARC format to link resolution providers.

The effectiveness of DOIs and OpenURLs has been well recognized. Using DOIs for e-books is a trend and sending title lists to link services providers KB is a standard practice for Western language resources providers. However, most Chinese providers do not provide DOIs or OpenURLs and are reluctant to communicate with link services providers directly for various reasons, such as security concerns, short staffing, or metadata creation cost. As a result, approximately 85% of the 35,000 UC-subscribed Chinese journals have either bad quality or no metadata in the SFX KB. This means that these 30,000 titles can hardly be discovered and accessed via SFX OpenURL service if users link from a non-traditional search utility such as Google Books. Similar complaints were heard from other subscribing institutions in Hong Kong and North America. China National Knowledge Infrastructure (CNKI) is the only UC vendor in China that has been providing title lists to Ex Libris via its agent. Through numerous communications, the CNKI leader has promised to follow KBART's recommendation in creating separate entries for title changes which will resolve most of CNKI's linking issues in SFX.

Bide (2015) pointed out that cost may be a barrier implementing DOI, "but all too often, people focus on direct, identifiable cost, and give little thought to benefit" ("Challenges," para 12). Given that the KBART title list is not costly to produce and some Chinese vendors like CNKI and Wanfang Data have been providing publishers free DOI registration for article-level objects, cost should not be the excuse for not providing DOIs and quality metadata to KBs.

KBART: A UNIVERSAL EXCHANGE FORMAT FOR TITLE LEVEL METADATA & LIBRARY HOLDINGS

KBART addresses issues related to the data supplied to knowledge bases and improves efficiency and effectiveness of linking. OpenURL connects users to library subscribed resources, (see section 2.2 and Figure 7). A significant part of this access flows through OpenURL link resolvers, which relies on detailed knowledge base holdings data provided by the content publishers. It is therefore vital that the metadata in knowledge base is accurate and up-to-date (NISO, 2014a). Illustrated in Figure 8 is the data flow through the supply chain (NISO/UKSG, 2010). If title level metadata in knowledge bases is inaccurate, it would undermine the purpose of OpenURL, and as a result, impact on electronic resource discovery and access.

KBART Phase I Recommended Practice
Figure 2: Flow of data through the supply chain
The author supplies manuscripts to the publisher (1). The publisher sends final article full text and associated metadata to one or more online delivery services (content hosts, aggregated full text databases, archive sites, etc.) (2). The content host is frequently responsible for delivering the article metadata to other discovery tools, such as A&I (abstracting and indexing) databases, subscription agent gateways, and search engines (3). Content hosts and full-text databases also deliver holdings information and sometimes MARC records to knowledge base developers (4) or directly to the institution (5) in order to inform those systems about the extent of content on the delivery platform. Some knowledge base services will then send holdings data to institutional A-to-Z lists and link resolvers (6); the institution uses this information to configure its own collection. Knowledge base developers may also send MARC records to the institution for inclusion in the institutional OPAC (7). In this scenario, the institution must supply its knowledge base developer with details of its holdings for content such as e-journals, etc. (8).

Figure 8. KBART flowchart of data transferring through the supply chains. Reprinted from http://www.uksg.org/sites/uksg.org/files/KBART_figure2.jpg

KBART Phase I working group identified some common issues related to poor metadata that prevents OpenURLs from linking to appropriate contents. These issues are: identifier inconsistencies (see Figures 9-10, Lam, 2014), title inconsistencies (see Figure 11-12, Lam, 2014), incorrect data coverage, inconsistent data formatting, and inconsistencies in content coverage description (see Figure 13, Lam, 2014), embargo inconsistencies, data format and exchange, outdated holdings data, and lack of customization (NISO, 2010a).

Figure 9. Same ISBN for different editions. Reprinted from Chinese e-resource metadata problems that cause access issues [PowerPoint slides], by C. Lam, 2014, slide7. Reprinted with permission.

Figure 10. Same ISBN for different titles by the same publishers. Reprinted from Chinese e-resource metadata problems that cause access issues [PowerPoint slides], by C. Lam, 2014, slide8. Reprinted with permission.

Figure 11. Title inconsistencies: missing part title. Reprinted from Chinese e-resource metadata problems that cause access issues [PowerPoint slides], by C. Lam, 2014, slide10. Reprinted with permission.

Figure 12. Title inconsistencies: recorded title incorrect. Reprinted from Chinese e-resource metadata problems that cause access issues [PowerPoint slides], by C. Lam, 2014, slide13. Reprinted with permission.

Figure 13. Inconsistencies in content coverage description. Reprinted from Chinese e-resource metadata problems that cause access issues [PowerPoint slides], by C. Lam, 2014, slide35&36. Reprinted with permission.

These examples illustrate how inaccurate metadata affects OpenURL linking. The need for standardized practice for data exchange through the supply chain becomes imperative. KBART provides a solution of using a simple unified metadata exchange format. This data exchange format includes a common set of metadata fields for transmission of holdings metadata from content providers to link resolver knowledge bases. The work of Phase I focused mostly on metadata for journals, and the work of Phase II expanded KBART to include metadata for consortia, open access, e-books, and conference proceedings.

To help improve the supply of accurate and up-to-date metadata to link resolvers, all stakeholders in the supply chain, particularly Chinese content providers, are encouraged to endorse the KBART Phase II Recommended Practice, use KBART file template, and file name convention. Librarians can use the KBART file template as the title level metadata checklist to ensure and validate files provided by contents providers:

- Comply with KBART recommended practices, such as metadata elements, formats, the sample file (see Figure 14, NISO, 2014a)

Sample1a: E-books only

Note: The spreadsheet for this data has been split into three sections for viewability in this document. The second and third sections represent additional columns that were appended to the first section in the original spreadsheet.

publication_title	print_identifier	online_identifier	date_first_issue_online	num_first_vol_online	num_first_issue_online	date_last_issue_online	num_last_vol_online	num_last_issue_online
Dante and Aquinas	978-1-909188-03-7	978-1-909188-07-5	2013-05-15					

title_url	first_author	title_id	embargo_info	coverage_depth	notes	publisher_name	publication_type
http://www.ubiquitypress.com/files/DanteAndAquinas.pdf	Ryan	10.5334/bad		fulltext	Published under CC-BY license	Ubiquity Press	monograph

date_monograph_published_print	date_monograph_published_online	monograph_volume	monograph_edition	first_editor	parent_publication_title_id	preceding_publication_title_id	access_type
2013-05-15	2013-05-24		1	Took			F

Figure 14. KBART sample file for e-books. Reprinted from *Knowledge Bases and Related Tools (KBART) Recommended Practice: A Recommended Practice of the National Information Standards Organization* (p. 28), by the KBART Phase II Working Group, 2014, Baltimore, MD: National Information Standards Organization (NISO).

- Supply the most comprehensive, accurate, and timely metadata to link resolution services

- Follow the order of data fields strictly

- Add any extra locally added data fields at the end

- Record every title, including former, later, or split titles in a separate entry

- Use the recommended structure and meaningful file names.

The benefits of following KBART recommended practice are:

- Vendor provides content metadata to knowledge bases in a universally accepted data format

- The standardized KBART file makes file transfer and knowledge base updates easier to handle

- Optimizes process and fosters smoother interaction between all stakeholders of knowledge bases supply chain

- Knowledge bases and library users benefit from provision of higher-quality data supplied by content providers

- Publishers expose content to greater usage by accurate linking to their contents

- Library has immediate return on investment when resources discovery and access become more efficient and effective

To help stakeholders in the supply chain for Chinese electronic resources, CEAL ERMB members work collaboratively with CALIS[5]/DRAA[6] on translating KBART Recommended Practice and conduct research on KBART application and its benefits and limitations on the efficiency and effectiveness of OpenURL linking for both Chinese and Western language resources.

In a recent article, Pesch (2014) provided some cases to illustrate that in addition to improving OpenURL linking through the definition of best practices for exchanging title-level metadata, KBART continues to gain traction and has the potential to be an initiative that improves not only link resolver quality but also improves discovery services, electronic resource management (ERM) systems, and even Counting Online Usage of Networked Electronic Resources (COUNTER) usage reports.

ODI: INDEXED DISCOVERY SERVICES
FOR CONTENT COVERAGE AND EXCHANGE

Several major discovery products, based on centrally indexed searches have been released to the market since early 2009 and have been widely deployed in libraries globally (NISO, 2014b).

The leading products are EBSCO Discovery Service (EDS), Ex Libris' Primo, Serial Solutions' Summon, and OCLC's WorldCat Local. There is also the SuperStar Chinese Discovery System developed in China. These services use a pre-aggregated central index to enable searching across a wide range of library related resources, both licensed and free, from multiple providers. They also offer more sophisticated capabilities and faster performance than those provided by systems relying on federated search technologies (NISO, 2014b).

However, there are issues all stakeholders are facing, discussed by both Kelley and members of the ODI Working Group in 2012 (Kelley, 2012). First, librarians, content providers, and discovery services providers often talk across one another when describing their respective needs and offerings, and no common vocabulary is used to describe what they do. Second, librarians expect that the large central index that underpins discovery services will maximize awareness and usage of library's entire collection, particularly for electronic subscriptions—which in some libraries accounts for 85% of material budget. But lack of visibility for what content is included in the central index, and what content had been indexed using metadata, abstract, full text, and/or subject headings makes it difficult for librarians to measure the impact after implementation of such services. Third, the indexed search is still governed by proprietary deals between discovery providers and information providers, which results in a blurry and inconsistent "ecosystem" that underserves libraries and users. Fourth, differences in access between indexing only metadata versus full text, as well as the depth and frequency of indexing, can make a dramatic difference in results when selecting a discovery services product. Lack of transparency in each discovery service's coverage of its index makes it hard to determine which product may offer the best coverage for a given library's subscriptions. Fifth, timely, standards-compliant usage and referral statistics were high on the wish list of content providers and librarians. The discovery vendors do provide usage statistics, but the formats vary from vendor to vendor and may not give certain content providers adequate credit for their contributions. Sixth, the transparency must extend to relevance ranking, which determines what results will float to the top and is a critical element in a given service's ultimate success or failure. Seventh, A&I providers are concerned about their position in discovery landscape. Although all the discovery providers say they fully appreciate how access to A&I databases is critical for deep and serious research and a key differentiator for libraries in their struggles to compete against search engines like Google, many A&I providers are wary when it comes to the discovery services and decline to add their bibliographic databases to the services, fearing doing so could threaten their lifeblood.

All of these issues brought stakeholders together with a growing need for best practices. They worked with NISO and launched the Open Discovery Initiative in 2012. After surveying the community, the recommended practices were released in 2014. The list provides recommendations for data exchange including data formats, method of delivery, usage reporting, frequency of updates, and rights of use. It offers a way to libraries to assess content providers' participation in discovery services. It is a model by which content providers work with discovery service vendors via fair and unbiased indexing and linking (NISO, 2014c). An important element of the ODI recommended practice is the disclosure to libraries of which metadata elements are included in discovery services. Conformance checklists to be used by discovery services providers and content providers are included as appendices to the Recommended Practices (Lagace, 2014).

The recommended practices for content providers (Section 3.2) are:

- Participation: provide core metadata and full-text/original content as well as enriched content
    - Core metadata elements – basic citation metadata (author, title, publisher, date, type, format, etc.)
    - Enriched content – indexing data (A&I data like subject headings), full text or transcript, abstracts/description
- Disclosure: provide information to libraries related to level of participation
- Technical formats: use existing standards to facilitate data exchange (NISO, 2016).

The recommended practices for discovery service providers (Section 3.3) are:

- Disclosure: provide key information in a consistent, usable form to libraries about content indexed to facilitate evaluation

- Linking: linking and relevancy methods should not introduce bias to particular content providers; libraries should determine linking choices; annual disclosure related to neutrality

- Data transfer: use existing protocols and provide documentation, preferences, and indication of impact on different processes to content providers

- Usage statistics for content providers (Section 3.3.4): provide simple and frequent statistics:

    - Total number of searches

    - Result clicks

    - Total number of click-throughs

- Usage statistics for libraries (Section 3.3.4): provide simple and frequent statistics:

    - Total number of searches per month

    - Total number of unique visitors per month

    - Total number of click-throughs per month

    - Top 500 search queries for the last period

    - Top 100 referring URLs to the discovery service for the last period (NISO, 2016).

Content providers and discovery service providers are encouraged to follow the recommended practices and investigate potential changes in their contributions to library services via indexed delivery. This will help optimize and simplify the process of data exchange between participating discovery vendors and content providers, ensure participating discovery vendors are following fair and unbiased indexing and linking practices, mitigate technical and legal issues that might hinder broader participation by content providers or potential discovery service creators, and allow libraries to understand how their licensed content is included in discovery systems (NISO, 2016).

COMPLYING WITH STANDARDS THROUGH EDUCATION AND LICENSING

To tackle issues related to discovering and accessing Chinese, Japanese and Korean electronic resources, The Council on East Asian Libraries (CEAL) established the Task Force on Metadata Standard and Best Practice (CEAL ERMB) in November 2013. The ERMB conducted a survey between January and February 2014, to investigate the awareness of standards for content organization of digital resources. The survey was sent to both Chinese vendors and Chinese studies librarians. The findings show that the top five standards that vendors follow are MARC 21, other classifications, ISSN, ISBN and other publication identifier, PIE-J and Library of Congress Subject Headings. It shows that the major standards, KBART, DOI and OpenURL, have not yet gained enough attention. Vendors and librarians were not clear on some of the key issues: they believe that metadata standards do not accommodate the needs of CJK resources/scripts, and that metadata standards do not necessarily eliminate problems. Findings also show that both vendors and librarians lack metadata expertise to provide required metadata and implement certain metadata standards.

One of the strategies for implementing NISO metadata standards is through training and education. Various organizations have actively engaged in this area at the national and local level. In 2014, CEAL ERMB conducted the *Workshop on Electronic Resources Standards and Best Practices: What Do Bibliographers, Catalogers, Publishers*, *and Vendors Need to Know in Philadelphia*. Over 70 librarians and vendors attended the workshop and this was the first time that NISO standards were presented to Chinese vendors. As the key driver for information industry standards, NISO has offered intensive training programs for a number of years. The NISO Education Committee was first formed in early 2007 in response to the training need from its broad constituency. The Committee determines topics, sets agendas, and recruits speakers. Twenty-five webinars or virtual conferences have been scheduled for 2016 and past training events can also be viewed on NISO's website (http://www.niso.org/news/events/2016/). NISO staff are frequently asked to make presentations at various conferences, workshops, and similar events and slide presentations are available for viewing on the NISO website (http://www.niso.org/news/on-the-road/). While librarians have benefited from NISO's efforts, more efforts are needed in particular for Chinese vendors and the Chinese librarian community. Only active engagement in learning and implementing NISO standards would help access to Chinese digital resources and further realize the value of those resources in teaching, research and pleasure reading.

Another strategy for enforcing the implementation of metadata standards is regulating it through the licensing of digital resources. Over the years, licensing terms have been updated to respond to the development of technology and pricing models. NISO standards have been added to several major model licenses in recent years. As a consortium of over 250 academic and independent research libraries in the U.S., Canada and Hong Kong, the Center for Research Libraries has developed a model license to serve as a statement by the academic library community of what it considers acceptable policy and practice for licensing digital information. This model license includes a clause on discovery of licensed materials, stating that vendors should "provide discovery service vendors on an ongoing basis the citation and complete descriptive metadata (including all subject headings, abstracts, and keywords), and full-text content necessary to facilitate optimal discovery and accessibility of the content" (Center for Research Libraries, 2014). Vendors should "comply with the most current version of the OpenURL standard (ANSI/NISO Z39.88), and will provide a mechanism for persistent links to content" (Center for Research Libraries, 2014). The model license specifies that vendors should provide customers "prior to the beginning of the calendar year within the current term, an itemized holdings report that specifies the titles included in the Licensed Materials for the next subscription term…If the Licensed Materials include content covered by the Knowledge Bases And Related Tools (KBART) Recommended Practice, itemized holdings lists for the Licensee will be reported in KBART-compliant format" (Center for Research Libraries, 2014).

The model license recommended by the Canadian Research Knowledge Network includes a clause that the publisher should "provide to the Consortium before December 31 of each year within the subscription period, in KBART-compliant format, an itemized holdings report that specifies the Licensed Materials accessible to the Members for the upcoming calendar year" (Canadian Research Knowledge Network, n.d.). The license also states that a publisher should use "reasonable and timely efforts to provide to third party vendors of Discovery Services, on an ongoing basis, with as comprehensive content for indexing as possible, including citation metadata (including subject headings and keywords), abstracts and full-text to facilitate optimal discovery of the content for the benefit of Authorized Users." (Canadian Research Knowledge Network, n.d.).

The model license recommended by the California Digital Library (CDL) includes a clause stating that, "The University of California is committed to the use of the emerging OpenURL standard to allow linking to related materials in other locations. If Licensor does not use the OpenURL standard, Licensor staff will provide information to Licensee upon request to assist the Licensee in

creating links directly from UC's library catalogs and licensed resources to the content at the journal, issue and article levels" (California Digital Library, 2014). CDL Model License does not include languages about other NISO standards, however, it recommends that librarians use *Technical Requirements for Licensed Resources* (California Digital Library, 2015) when negotiating and acquiring digital resources. This document requires that vendors provide a full title list as compliant with the KBART standard for title or item-level content. For construction of inbound URLs, vendors should follow NISO OpenURL standard Z39.88-2004, and the elements used in constructing the URL must be common citation data elements and cannot be tied to an identifier specific to the vendor such as a unique article ID number. To facilitate discovery services, the CDL requests that vendors make their records available for loading into web-based discovery services, so that UC patrons have centralized access to a broad range of materials. Sharing of records should follow the NISO ODI and NISO metadata and best practices.

Including languages concerning NISO standards in license, if implemented strictly, would help alleviate access and discovery problems with Chinese digital resources. What has happened, often times, is that when acquiring digital resources, technical standards tend to be compromised or ignored due to negotiating for better pricing and service terms. The direct result is a negative impact on access and discovery of acquired resources, which harms vendors/publishers, libraries, and ultimately, library users. As major players in the chain, vendors and librarians need to work together to improve the situation and to improve the usage of Chinese digital resources overseas, as they are one of the major materials sources for faculty, students and researchers.

CONCLUSION

NISO standards and recommended practices offer solutions to most of the content presentation and metadata issues with Chinese digital resources; however, these issues can only be addressed and resolved collectively by all players in the supply chain. Promotion of the compliance with standards also requires collaboration among the stakeholders. To facilitate such cooperation and collaboration, developing international standards that incorporate needs from all language resources is critical and is also our ultimate goal.

CEAL ERMB members and the CALIS/DRAA Standards and Recommended Practices Research Task Group co-launched the *China-U.S. Translation and Research Collaboration Project on Electronic Resource Standards and Recommended Practices in the United States* in April 2016. This project aims to develop national and international standards for managing Chinese digital resources, and to communicate standards and compliance with vendors and/or information providers of East Asian digital publications.

Implementing metadata standards also requires training and education, the enforcement of the implementation through the process of licensing of digital resources, and the fostering of trust between vendors and libraries. Bide (2015) made the important point that standards implementation is not simply technology, but trust. Trust between vendors and librarians is established on the foundation of mutual benefit.

References

Atwill, Y. Y. (2005, November). E-journals from China: Technical and collection issues. *Journal of Academic Librarianship, 31*(6), 598-604. doi:10.1016/j.acalib.2005.09.001

Beam, J. T., & Copeland, N. S. (2001). Electronic resources in union catalogs: URLs and accessibility issues. *Serials Review, 27*(3/4), 33-45. Retrieved March 16, 2016, from http://search.ebscohost.com/login.aspx?direct=true&db=lxh&AN=6106618&site=ehost-live

Bide, M. (2015). The DOI, twenty years on. *D-Lib Magazine, 21*(7/8). http://doi.org/10.1045/july2015-bide

California Digital Library. (2014). *Standard License Agreement.* Retrieved March 1, 2016, from California Digital Library: http://www.cdlib.org/services/collections/docs/toolkit/Model_License_LATEST_Revised_03-2014.docx

California Digital Library. (2015). *Technical Requirements for Licensed Resources.* Retrieved March 1, 2016, from California Digital Library: http://www.cdlib.org/gateways/vendors/guidelines_technical.html

Canadian Research Knowledge Network. (n.d.). *Model License Agreement.* Retrieved March 1, 2016, from Canadian Research Knowledge Network: http://crkn.ca/programs/model-license

Center for Research Libraries. (2014). *Liblicense Model License Agreement with Commentary.* Retrieved March 1, 2016, from Center for Research Libraries: Center for Research Libraries (2014), Liblicense Modhttp://liblicense.crl.edu/wp-content/uploads/2015/05/modellicensenew2014revmay2015.pdf

Corporation for National Research Initiatives. (n.d.). *Handle.Net® registry*, March 2, 2016. Retrieved March 20, 2016, from Handle.Net® registry: http://www.handle.net/index.html

Crossref. (n.d.). *FastFacts* Retrieved November 29, 2016, from http://www.crossref.org/01company/16fastfacts.html Ferguson, A. W., & Ko, A. (2004, November). eBooks in China. *Against the Grain, 16*(5), 42-50. doi:10.7771/2380-176X.4066

Gul, S., Mahajan, I., & Ali, A. (2014, September). The Growth and decay of URLs citation: a case of an online library & information science journal. *Malaysian Journal of Library & Information Science, 19*(3), 27-39. Retrieved April 18, 2016, from http://web.b.ebscohost.com/ehost/pdfviewer/pdfviewer?vid=5&sid=87ffe589-a624-4926-9dc9-7a2b343bb34f%40sessionmgr120&hid=124

International DOI Foundation. (2016). *Resolution*, February 22, 2016. doi:http://dx.doi.org/10.1000/182

International Standard Serial Number International Center. (2015, January). *ISSN Manual.* (I. S. Center, Ed.) Retrieved September 25, 2016, from ISSN: http://www.issn.org/wp-content/uploads/2013/09/ISSNManual_ENG2015_23-01-2015.pdf

ISO. (2012, May 10). *Digital object identifier (DOI) becomes an ISO standard*. (E. Gasiorowski-Denis, Editor) Retrieved March 20, 2016, from International Organization for Standardization News: http://www.iso.org/iso/home/news_index/news_archive/news.htm?refid=Ref1561

ISSN Zhongguo guo jia zhong xin. (n.d.). *Cha xun ISSN* [Search ISSN portal]. Retrieved March 20, 2016, from ISSN Zhongguo guo jia zhong xin: http://www.nlc.gov.cn/newissn/cxissn/

ISSN China Center. (n.d.). *ISSN Application*. Retrieved March 20, 2016, from ISSN Zhongguo guo jia zhong xin: http://www.nlc.gov.cn/newen/fl/issncc/issnapp/

Kasprowski, R. (2012). NISO's IOTA Initiative: Measuring the quality of OpenURL links. *The Serials Librarian, 62*(1-4), 95–102. doi:10.1080/0361526X.2012.652480

Kelley, M. (2012, October 15). Coming into focus: web-scale discovery services face growing need for best practices. *Library Journal, 137*(17), 34-40.

Kumar, B. T., & Manoi, K. S. (2012). Persistence and half-life of URL citations cited in LIS open access journals. *Aslib Proceedings, 64*(4), 405-422. doi:http://dx.doi.org/10.1108/00012131211244752

Lagace, N. (2014, December). NISO releases recommendations from the Open Discovery Initiative: Promoting transparency in Discovery Working Group. *Serials Review, 40*(4), 287-288. doi:DOI: 10.1080/00987913.2014.978244

Lagace, N., Kaplan, L., & Leffler, J. J. (2015). Actions and updates on the standards and best practices front. *The Serials Librarian, 68*(1-4), 191-196. doi:10.1080/0361526X.2015.1017420

Lam, C. (2014, March 15). *Chinese e-resource metadata problems that cause access issues.* Retrieved March 16, 2016, from CEAL Committee on Technical Processing: http://www.eastasianlib.org/ctp/Workshops/2014/CEAL_ERMB_Connie_rev.pptx

Ma, B.-H. (2012, March 14). *Strengthening the Chinese electronic resources supply chain with standards and best practices.* Retrieved March 16, 2016, from CEAL Committee on Technical Processing: http://www.eastasianlib.org/ctp/Meetings/2012/Ma_StrengtheningChineseE.pptx

Marshall, D., & Reynolds, R. (2013). Having e-Journal title and ISSN problems? Have some PIE-J! *The Serials Librarian, 65*(1), 63–68. doi:10.1080/0361526X.2013.800465

National Information Standards Organization. (2010a, January). *KBART: Knowledge Base and Related Tools.* Retrieved March 16, 2016, from UKSG: http://www.uksg.org/sites/uksg.org/files/KBART_Phase_I_Recommended_Practice.pdf

National Information Standards Organization. (2010b). The OpenURL Framework for Context-Sensitive Services. Baltimore, MD, U.S.: NISO Press. Retrieved March 21, 2016, from http://www.niso.org/apps/group_public/download.php/14833/z39_88_2004_r2010.pdf

National Information Standards Organization. (2010c). *Syntax for the digital object identifier.* (S. f. identifier, Ed.) Retrieved March 20, 2016, from National Information Standards Organization: http://www.niso.org/apps/group_public/download.php/14689/z39-84-2005_r2010.pdf

National Information Standards Organization. (2013). *PIE-J: The Presentation & Identification of E-Journals.* Retrieved March 20, 2016, from National Information Standards Organization: http://www.niso.org/apps/group_public/download.php/10368/rp-16-2013_pie-j.pdf

National Information Standards Organization. (2014a, June). *Knowledge Bases and Related Tools (KBART) Recommended Practice.* Retrieved March 16, 2016, from National Information Standards Organization: http://www.niso.org/publications/rp/rp-9-2014/

National Information Standards Organization. (2014b). *ODI.* Retrieved March 26, 2016, from National Information Standards Organization: http://www.niso.org/workrooms/odi/

National Information Standards Organization. (2014c, June 25). *Open Discovery Initiative: Promoting transparency in discovery.* Retrieved March 26, 2016, from National Information Standards Organization: http://www.niso.org/workrooms/odi/publications/rp/rp-19-2014

National Information Standards Organization. (2016). *NISO Open Discovery Initiative Brochure.* Retrieved March 26, 2016, from National Information Standards Orgaznization: http://www.niso.org/apps/group_public/download.php/15232/NISO%20ODI%20brochure.pdf

NISO/UKSG KBART Working Group. (2010, January). *KBART Phase I Recommended Practice: Figure 2: Flow of data through the supply chain.* Retrieved March 16, 2016, from UKSG: http://www.uksg.org/sites/uksg.org/files/KBART_figure2.jpg

Ouyang, J. (2013). Shi zi tu shu guan xin xi zi yuan de ke fa xian xing yan jiu [Research on the Discoverability of Digital Library Information Resources]. *Tu shu guan lun tan , 33*(1), 32-37.

Pesch, O. (2014). The KBART's potential beyond OpenURL linking. *The Serials Librarian*(67), 231-239. doi:10.1080/0361526X.2014.960643

Prithviraj, K. R., & Kumar, B. T. (2014, March). Corrosion of URLs: Implications for electronic publishing. *IFLA Journal, 40*(1), 35-47. doi:http://dx.doi.org/10.1177/0340035214526529

Sa, L., Wen, Q., & Qu, Y. (2011). Shu zi zi yuan fa xian fu wu fa zhan qu shi yan jiu [Study on the Developing Trend of Digital Resources Discovery Service]. *Xian dai qing bao, 31*(9), 75-79. doi:10.3969/j.issn.1008-0821.2011.09.018

Sun, Y. (2013). 2013 nian tu shu guan qian yan ji shu lun tan (IT4L): "zi yuan fa xian zhi lv" yan tao hui zong su [2013 Information Technology for Library (IT4L: Summary of discovering the ultimate beauty of the Library Services]. *Shu zi tu shu guan lun tan*(7), 68-70. doi:10.3772/j.issn.1673—2286.2013.07.014

Xin hua yue du. (2014, November 17). *Di er jie tu shu guan xian dai ji shu xue shu yan tao hui zai guo jia tu shu guan ju xing [The 2nd Library Symposium on Advanced Technology was held in the National Library of China].* Retrieved September 9, 2016, from Xin hua wang: http://news.xinhuanet.com/book/2014-11/17/c_127219009.htm

Xue, S. (2014, March 15). *Highlights of the Survey on Metadata Standards and Best Practices for Chinese E-Resources.* Retrieved March 16, 2016, from CEAL Committee on Technical Processing: http://www.eastasianlib.org/ctp/Workshops/2014/Susan%20Highlights_Chinese.pptx

Zhai, H. (2014). Gao xiao tu shu guan shu zi zi yuan jian she de ce lue yan jiu - yi Anhui sheng wei li [Study on strategies for the development of digital resources in academic libraries - a case study on Anhui Province]. *Nong ye wang luo xin xi*(8), 89-91.

Zhang, H. (2014, November). Xue wei lun wen zhu ce DOI zhi xiao yi yu shi wu [The benefit and implementation of DOI registration for theses and disserations]. *Guo li Chenggong da xue tu shu guan guan kan*(23), 9-18. Retrieved September 12, 2016, from http://www.lib.ncku.edu.tw/journal/23/pdf/23_02.pdf

About the authors

Bie-Hwa Ma is Chinese Language Metadata Specialist at University of California, San Diego Library.

Shi Deng is the Head of CJK Cataloging & Metadata Unit at University of California, San Diego Library.

Susan Xue is the Head of Information and Public Services and Electronic Resources Librarian at University of California, Berkeley Library.