

# Making Metadata Maker

## A Web Application for Metadata Production

**Myung-Ja K. Han, Nicole E. Ream-Sotomayor,  
Patricia Lampron, Janet Weber, and Deren Kudeki**

*Cataloging and metadata operations in academic libraries are focusing on original cataloging of their unique and hidden collections that have not been available to users because of a lack of metadata. However, creating MARC format metadata is an expensive process; libraries need professional catalogers with appropriate experience and knowledge or must train staff to do the work. To improve the cataloging and metadata creation workflow, the University of Illinois at Urbana-Champaign Library developed a web application, Metadata Maker, which allows anyone to create metadata in four different formats, including MARC21 for an online public access catalog, regardless of their familiarity with metadata standards or systems that utilize the metadata. Released as an open source application, Metadata Maker supports diacritics and Unicode non-Roman language encoding, and creates metadata records that ensure discovery and access of unique library collections.*

As more resources purchased by libraries come with vendor provided cataloging records or via other libraries through cooperative cataloging initiatives, cataloging and metadata operations in academic libraries are focusing on processing more unique materials and hidden collections that have not been available to users because of a lack of metadata. To provide metadata, libraries generally employ professional catalogers with subject knowledge and appropriate cataloging experience to make these hidden and possibly valuable library collections searchable and discoverable in a timely manner. Recent budgetary issues and a shift in library priorities led to these positions being downsized or eliminated, and libraries must find alternatives to facilitate metadata creation.<sup>1</sup> At the University of Illinois at Urbana-Champaign (UIUC) Library, the loss of professional cataloging positions due to retirements and the increasing volume of resources that need catalog records has imposed a change in cataloging workflows that rely on temporary staff with little to no experience in metadata creation. While the UIUC Library has worked to train these staff to create metadata in MARC format and to use appropriate cataloging software, such as integrated library systems and shared cataloging systems (e.g., OCLC's Connexion), this training is an intensive and time-consuming process. Because of the temporary nature of many of these staff members, the professional catalogers are in a constant state of training new employees, monitoring their work, and providing appropriate feedback to protect

**Myung-Ja K. Han** (mhan3@illinois.edu) is a Metadata Librarian, **Nicole E. Ream-Sotomayor** (ream2@illinois.edu) is the Foreign Language Cataloging Coordinator, **Patricia Lampron** (lampron2@illinois.edu) is a Metadata Services Specialist, **Janet Weber** (janetw@illinois.edu) is Cataloging Services Coordinator, and **Deren Kudeki** (dkudeki@illinois.edu) is Project Programmer at the University of Illinois at Urbana-Champaign.

Manuscript submitted March 4, 2015; returned to authors May 13, 2015 for revision; revised manuscript submitted June 17, 2015; returned to authors for minor revision September 30, 2015; revised manuscript submitted October 9, 2015; accepted for publication November 9, 2015.

metadata quality. Another option that libraries may consider is the outsourcing of cataloging work to vendors, but that also incurs a substantial cost to accomplish the task.

To improve the productivity of their cataloging and metadata creation workflow, the UIUC Library sought to develop a web application, Metadata Maker, that would allow anyone to create metadata in various formats, regardless of their familiarity with metadata standards or systems that utilize the metadata.<sup>2</sup> With information that is readily available from the item in hand, a user can create quality metadata that ensures the discoverability of resources in the library's various asset management systems including the online public access catalog (OPAC). This paper discusses the emerging need for libraries to have a metadata creation tool that enables metadata creation in different formats, and shares a detailed description of the project development process and the initial user testing results.

### Literature Review

The need for an efficient metadata creation tool for backlogs of library materials is steadily increasing as library budgets and experienced cataloging positions are being reduced, and libraries are shifting priorities to electronic resources and digitization efforts. Boydston and Leysen studied the changing nature of the roles and responsibilities of the cataloging librarian by examining thirty-two responses to their survey on the topic.<sup>3</sup> While the study discusses in depth the shift in priorities and skills toward electronic resources and new descriptive standards including non-MARC metadata, it also reports that 60.71 percent of responders indicated that cataloging positions had been eliminated. They cite "budget, reorganization, and retirements" as top reasons for the elimination of these positions.<sup>4</sup> Ithaca S+R Library Survey 2013 results, with responses from 499 academic libraries, indicate that many library directors are experiencing constraints because of limited budgets, while their focus is shifting toward digital preservation and special collections, leaving the purchase and processing of print materials on the back burner.<sup>5</sup> The survey also shows that almost 30 percent of respondents predict a reduction of staff in "technical services, metadata, and cataloging."<sup>6</sup> Of the sixteen categories of "staff resources," the metadata and cataloging staff category was predicted to see the largest decrease.<sup>7</sup>

A decrease in experienced catalogers also brought a change in the roles and responsibilities of cataloging librarians. Sapon-White found that traditional cataloging work has shifted to "paraprofessional staff" and graduate students who required regular cataloging training by catalogers.<sup>8</sup> Sapon-White asserted that while the training program organized for the Oregon State Library was successful, it required one-on-one training and then weekly follow-up

training, which spanned a two year period, and only in one aspect of cataloging: subject analysis.<sup>9</sup>

Compounding the issue of staffing and resources for cataloging print materials is the prevalence of backlogged print materials among academic libraries. Jones discusses in depth the problem of hidden collections and inaccessible resources due to the lack of metadata.<sup>10</sup> Not only do unprocessed materials go undiscovered by library users, they also cause errors or duplication in acquisitions, can be lost or stolen, and lead to poor donor relations resulting from "not making collections available in a timely fashion."<sup>11</sup> Citing the results of the 1998 ARL Survey by Panitch, the white paper states that "15 percent of collections on average remained unprocessed or uncataloged" and suggests that libraries must develop policies and workflows for levels of access, while taking into account characteristics of the individual item or collection.<sup>12</sup>

Metadata Maker is designed to allow staff with minimal training in the intricacies of cataloging and metadata standards, but possibly with more subject or language knowledge, to create metadata describing these backlogged materials to provide adequate access to both library users (for research) and staff (for resource maintenance). In its attempt to provide access to library materials in a timely manner, the UIUC Library's Metadata Maker project team had to determine the minimal standard of metadata quality acceptable for bibliographic metadata produced by the application. This minimal set of information was guided by the main goal of the project: to produce bibliographic metadata that facilitates discovery and access of the library resource both by users and librarians who use and manage them.

There is great debate regarding what constitutes quality bibliographic metadata across the library community. However, there is a general consensus that accuracy is important, plus the presence of appropriate access points and subject headings and the usefulness of the record in terms of search and retrieval. Snow reported on academic library catalogers' perceptions of quality catalog records from her study.<sup>13</sup> Referring to interviews and questionnaires completed by cataloging librarians, Snow described aspects of cataloging that determine quality. For the purposes of the study, "quality cataloging" is broken into four categories: technical details of the bibliographic record, adherence to standards, the cataloging process/workflow/staff, and impact on users/accessibility.<sup>14</sup> Of these four, technical details of the bibliographic record were discussed the most when describing quality cataloging, though almost 80 percent of respondents described aspects that represent two or more categories.<sup>15</sup> When asked to rank MARC data fields and subfields, the catalogers surveyed chose the top three data fields as 245\$a (Title Proper), 100 (Personal Name), and 650 (Topical Subject Heading), indicating that access points are perceived as

a component of a quality record.<sup>16</sup> Roy Tennant's visualization of MARC fields used on book records further supports Snow's findings, which shows that the 040 (Cataloging Source), 245 (Title Statement), 260 (Publication), 300 (Physical Description), and 650 (Subject Added Entry—Topical Term) are the most commonly appearing MARC data fields for such records.<sup>17</sup> Calhoun et al. addressed the concept of "quality data" by surveying both users and librarians on their expectations for search results.<sup>18</sup> One of the key findings of the study highlights librarians' emphasis on duplicate record merging and clean up, indicating the need for bibliographic metadata that have adequate information for matching resources to their proper record, which played an important role in the design of Metadata Maker.

### Project Development

The Content Access Management (CAM) unit of the Technical Services Division in the UIUC Library has gained valuable experience in creating metadata for digital resources and special collections during the last several years by implementing various information technologies, notably Extensible Markup Language (XML), Extensible Stylesheet Language for Transformations (XSLT), and simple programming languages including Python and JavaScript. Metadata creators for the Emblematica Online project, a digital portal for all digitized emblem books, are not trained in any metadata standards or systems.<sup>19</sup> An emblem book contains a collection of emblems, which are described as "a symbolic and often enigmatic image," and the image's accompanying text.<sup>20</sup> To provide access to a book and the emblems contained in it, metadata must be created for each emblem. Students and scholars who create emblem-level metadata use a Microsoft Excel spreadsheet as their input form, adding information about the individual emblem to designated rows and columns. Once the spreadsheet describing all the emblems in a book is complete, it is saved in XML Spreadsheet format. Using XSLT, the metadata in the XML Spreadsheet is then transformed to the SPINE metadata standard, a schema specially designed for describing emblem books and their emblems.<sup>21</sup> With the new workflow in place, the library does not need to conduct training on the metadata standard, and system access permissions are not required for each individual working on the project.

Relying on their previous experience with automated metadata creation workflows, a group within CAM, consisting of the Senior Metadata Librarian, the Foreign Language Cataloging Specialist, the Manager of Cataloging Services, and the Metadata Specialist, started the Metadata Maker project with support from an Innovation Grant from the University Librarian in October 2014 that allowed the group to hire a part-time research programmer for the project.

The project team sought to build a web-based application that allows anyone to create metadata in MARC21, MARCXML, Metadata Object Description Schema (MODS), and/or HTML page marked up with Schema.org semantics. While the project originally started as MARC Maker, the addition of non-MARC metadata output prompted a change in the project name. The application, currently in version 1.1, cues the user to record information about a resource, such as *Title*, *Name*, and *Keywords*, plus other bibliographic information, and provides options for the user to choose one or more of the output formats listed above. The data are then transformed into the selected format(s) and downloaded to the user's own computer. The data can subsequently be reviewed by another staff member with cataloging experience or a professional cataloger and ingested into the appropriate library system to allow for resource discovery.

### Metadata Design

The initial goal of the project was to create minimal-level bibliographic metadata in MARC21 format. Starting with the Library of Congress (LC)'s minimal-level record for books example and Snow's findings, the project team decided that the MARC record should include the data fields 1XX (Main Entry, if applicable), 245 (Title Statement), and keywords used to provide subject access.<sup>22</sup> In addition, data fields 260/264 (Publication, Distribution, etc. (Imprint)), and 300 (Physical Description) are added because this information is readily available from the item. For Metadata Maker to create rich metadata, each subfield is separated from its parent data field in application display so that users can add individual data attributes appropriate for each subfield into a separate element in the application. For example, in the Metadata Maker web form, the Title statement is divided into *Title* and *Subtitle*, and the Imprint statement is divided into *Place of publication*, *Name of publisher*, *Date of publication*, and *Copyright date*. This ultimately enables the application to encode information in each element within the proper MARC subfield with International Standard Bibliographic Description (ISBD) punctuation for the MARC21 format metadata. The application also provides sections to record other information easily gleaned from the item that is being cataloged, such as *Number of pages/volumes* (options are selected from a dropdown menu), *Language*, *ISBN*, and *Edition statement*. Also included in the application are questions that determine whether the item is a work of fiction or literature (*Is this item literature?*), and whether it includes illustrations (*Does the item include illustrations?*).

Because of the user testing findings, additional exploration of controlled vocabulary services, and the UIUC Library's cataloging needs, the elements available in Metadata Maker have changed since its initial release as version 1.0. In version 1.0, the application included names only for

**Table 1.** Elements Included in the Metadata Maker Application ((R) identifies an element as repeatable)

Element Name	Required/Optional
Title	Required
Subtitle	Optional
ISBN	Optional
Edition statement	Optional
Language	Required
Names (R)	Optional
Name of publisher	Optional
Place of publication	Optional
Country of publication	Optional
Date of publication	Optional
Copyright date	Optional
Number of pages/volumes	Optional
Dimensions	Required
Is this item literature? (If yes, choose from dropdown)	Optional
Does the item include illustrations?	Optional
Keywords (R)	Required
Note to the Cataloger	Optional

authors, and the element was labelled as *Author*. However, the user testing results revealed that testers added additional names available on the item into the *Note to the Cataloger* element. In version 1.1, the *Author* element label was changed to *Name*, and now users can select role information for the name from one of the six roles provided in the dropdown menu: artist, author, contributor, editor, illustrator, and translator. Depending on future needs, additional role values may be added to the dropdown menu. Also changed for version 1.1 is the implementation of Faceted Application of Subject Terminology (FAST) headings in the *Keywords* element, which are mapped to 6XX data fields accordingly. By facilitating FAST headings, users can now choose keywords from the already established controlled subject headings in addition to using any uncontrolled terms. Table 1 shows all available elements a user can input in Metadata Maker.

The application automatically adds default information into the metadata in MARC21 and MARCXML formats. For example, because the application was developed for monographic materials, MARC data fields 336 (Content Type), 337 (Media Type), and 338 (Carrier Type) are populated with the appropriate designators for a book and are added automatically during the transformation process. Data field 040 (Cataloging Source), Leader position 07 (Bibliographic Level), and Leader position 18 (Descriptive Cataloging Form) are also added automatically (see table 2).

In the end, the application allows users to create a metadata record that is close to a full-level record if all information is available within the item and recorded accordingly.

### Web Application Design

The application is presented as a web form (see figure 1) comprising the elements shown in table 1. There is a question mark icon next to each element that provides a short description with an example when the cursor hovers over it.

Among the sixteen elements, *Title*, *Language*, *Dimensions*, and *Keywords* are required. *Number of pages/volumes* requires a number to be entered under “pages” or “volumes,” or the “unpaged” box to be checked, as shown in figure 1. The remaining eleven elements are optional. Though Metadata Maker can record a range of descriptive elements, to make the majority of these elements required would exclude metadata creation for a great deal of material. As the goal of Metadata Maker is to facilitate a more efficient workflow for metadata creation, the project team sought to make the tool flexible enough to work with any materials, and so only the four elements mentioned above are required. Two elements, *Name* and *Keywords* are repeatable—a new box is added when the “+” sign located after the first text box is clicked. All other elements have a check box named “unlisted” beneath the text box, which the user can check when information for that element is not available in the item. During transformation into metadata, if an element is not populated, and “unlisted” is not checked, a message will appear as part of the validation process, as a measure of verifying that the information is not included in the resource. Values for three elements were designed to be chosen from the element’s dropdown menu: *Language*, *Country of publication*, and *Type of literature* to ensure consistent and quality metadata. Values for *Language* and *Country of publication* are based on controlled vocabularies available from LC.<sup>23</sup> The dropdown menu for *Number of pages/volumes* allows the user to record either the number of *pages* (for a single volume) or *volumes* (for multivolume sets). The *Number of pages/volumes* element also includes an “unpaged” box, which the user can check if the piece lacks page numbers.

A *Note to the Cataloger* field is provided to allow input of any additional information that should be included in the metadata or requires the cataloger’s attention. Information added in this element is initially transformed into the data field 500 (General Note) as a note for the cataloger who will review the metadata in OCLC Connexion. The decision to use the data field 500 instead of the 590 (Local Notes) was based on local practice and workflow. It was noted during testing that the *Note to the Cataloger* field was used for various types of information, such as presence of multiple languages, translations, or series, which may belong either in a public note or other MARC data fields. Since all catalog

**Table 2.** Mapping from the elements in the application to MARC, MODS, and HTML page marked up with Schema.org semantics (Rows in grey are added into MARC format metadata during transformation process.)

Element in Application	MARC	MODS	Schema.org
Title	245 1? ‡a	<titleInfo> <title>	name (Thing)
Subtitle	245 1? ‡b	<titleInfo> <subTitle>	name (Thing)
ISBN	020 __ ‡a	<identifier type="isbn">	isbn (Book)
Edition statement	250 __ ‡a	<originInfo> <edition>	
Language	008/35-37	<language> <languageTerm type="code" authority="iso639-2b">	inLanguage (CreativeWork)
Name (author)	100 1_ ‡a	<name type="personal"> <namePart type="given"> <namePart type="family">	Author
	100 1_ ‡e	<role> <roleTerm type="text" authority="marcrelator">author</roleTerm> <roleTerm type="code" authority="marcrelator">aut</roleTerm>	
	100 1_ ‡4	<name type="personal"> <namePart type="given"> <namePart type="family">	
Other Names	700 1_ ‡a	<role> <roleTerm type="text" authority="marcrelator">role term</roleTerm>	
role="editor"	700 1_ ‡e	<roleTerm type="code" authority="marcrelator">role code</roleTerm>	editor
role="contributor"			contributor
role="illustrator"	700 1_ ‡4		illustrator
role="translator"			translator
Name of publisher	264 _1 ‡b	<originInfo> <publisher>	publisher (CreativeWork)
Place of publication	264 _1 ‡a	<originInfo> <place type="text">	n/a
Country of publication	008/15-17	<originInfo> <place type="code" authority="marccountry">	n/a
Date of publication	008/06 008/07-10 264 _1 ‡c	<originInfo> <dateIssued>	datePublished (CreativeWork)
Copyright date	008/06 008/11-14 264 _4 ‡c	<originInfo> <copyrightDate>	copyrightYear (CreativeWork)
Number of Pages/volumes	300 __ ‡a	<physicalDescription> <extent>	numberOfPages (Book)
Dimensions	300 __ ‡c	<physicalDescription> <extent>	
Is this item literature?	008/33	<genre>	genre (CreativeWork)
Does the item include illustrations?	008/18 300 __ ‡b	<physicalDescription> <note>	
Keywords	653 __ ‡a	<subject> <topic>	keywords (CreativeWork)
Content Type	336 __ ‡a, b, 2	n/a	n/a
Media Type	337 __ ‡a, b, 2	n/a	n/a
Carrier Type	338 __ ‡a, b, 2	n/a	n/a
Cataloging Source	040 __ ‡a, b, e, c	n/a	n/a
Descriptive Cataloging Form	Leader 18	n/a	n/a

**Figure 1.** User interface of Metadata Maker, a web application for a metadata creation, with illustrations of required elements, drop down menu options, and FAST suggested terms.

records created by the application will be reviewed by a cataloger, using the data field 500 for all notes works for UIUC's local purposes and makes the application more user friendly for non-catalogers. The MARC field designation can be changed depending on each institution's implementation plan and practice; however, the *Note to the Cataloger* element should be used only when a professional cataloger will review metadata after non-catalogers create the metadata.

After all the available information is added, the user chooses one or more metadata output formats (MARC21, MARCXML, MODS, and/or HTML page marked up with Schema.org semantics), and the metadata downloaded to the user's computer. The application allows the user to name the file, or applies a default file name if one is not provided.<sup>24</sup> Figure 2 shows metadata in MARC format created by staff using Metadata Maker, and figure 3 shows the metadata after the professional cataloger enhanced it.

#### Metadata Formats Output

Because libraries work with many different metadata standards, the application currently creates metadata in four formats: MARC21, MARCXML, MODS, and HTML page marked up with Schema.org semantics. MARC21 was

```
000 00663cam a2200205ki 450
001 7757597
005 20150602104540.0
008 150602t20122012caua 000 u ful d
040 __ |a UIU |b eng |e rda |c UIU
100 1_ |a Kamara, Mamadou, |e author. |4 aut
245 10 |a Togue's alphabet.
264 _1 |a California : |b Friends of African Villages Libraries, |c [2012]
264 _4 |c ©2012
300 __ |a 38 pages : |b illustrations ; |c 21 cm
336 __ |a text |b txt |2 rdacontent
337 __ |a unmediated |b n |2 rdamedia
338 __ |a volume |b nc |2 rdacarrier
500 __ |a The item is in Pular.
650 _7 |a Pular dialect |2 fast |0 (OCoLC)fst01083771
650 _7 |a Alphabet books |2 fast |0 (OCoLC)fst00805964
```

**Figure 2.** MARC21 format record created by staff using the Metadata Maker.

```
000 00997cam a2200289ji 450
001 7226405
005 20150602100244.0
008 130509t20122012caua 000 0 ful d
035 __ |a (OCoLC)ocn899141392
040 __ |a UIU |b eng |e rda |c UIU |d UIU
050 _4 |a PL8184.Z9 |b P8553 2012
049 __ |a UIUU |o skip
100 1_ |a Kamara, Mamadou, |e author. |4 aut
245 10 |a Togue's alphabet / |c created by Mamadou Kamara with help from Togue.
264 _1 |a California : |b Friends of African Villages Libraries, |c [2012]
264 _4 |c ©2012
300 __ |a 38 pages : |b illustrations ; |c 21 cm
336 __ |a text |b txt |2 rdacontent
337 __ |a unmediated |b n |2 rdamedia
338 __ |a volume |b nc |2 rdacarrier
546 __ |a In Pular.
650 _7 |a Pular dialect |2 fast |0 (OCoLC)fst01083771
650 _7 |a Alphabet books |2 fast |0 (OCoLC)fst00805964
650 _0 |a Pular dialect |x Alphabet |v Juvenile literature.
650 _0 |a Pular dialect |x Vocabulary |v Juvenile literature.
655 _0 |a Alphabet books |z Africa.
994 __ |a C0 |b UIU
```

**Figure 3.** The same record after post processing—enhanced by a cataloger with a statement of responsibility, subject headings, classification, and others.

the first format chosen for output because the majority of metadata created locally from the application will be ingested into OCLC Connexion and Voyager, UIUC Library's integrated library system (ILS). It was decided that MARCXML, MODS, and HTML page marked up with Schema.org semantics should also be offered as output formats by the application. Currently, more metadata services require MARC records in XML format, and having MARCXML without the need to use another transformation process streamlines the workflow. The UIUC Library's digital preservation system requires MODS as its bibliographic metadata standard, so MODS was also selected as one of the options. The HTML page marked up with Schema.org

semantics was added as part of the UIUC library's ongoing contribution to Linked Open Data development work. The UIUC library has experimented with linked data by transforming 5.5 million bibliographic records and associated holdings data to linked data using Schema.org semantics that align with OCLC's linked data work.<sup>25</sup> For this experimentation, most MARC data fields and subfields were mapped to Schema.org semantics and transformed by Metadata Maker accordingly.

### Transforming Information to Metadata

To transform information submitted to the application into the four metadata formats, three different mappings were created, with one mapping used for both MARC21 and MARCXML (see table 2). For MARC21, since the metadata will be imported into OCLC Connexion, the output format for MARC21 metadata are an MRC (machine-readable) file. Transforming application data to metadata in MARC21 format is more challenging than other output metadata formats because some mappings utilize the information entered to determine the data in the fixed fields. However, because MARC21 and MARCXML have the same structure and the underlying schema is the same, transformation of these two metadata formats are done through similar code. Preparing transformation of information to MODS and HTML is simpler than MARC21 and MARCXML for two reasons: mapping to MODS and HTML do not require conditional mapping (for example, in a MARC record, the first indicator of the data field 245 is determined by the value in the *Names* element); and MODS and HTML are not affected by the white spaces in fixed field 008 and the leader in MARC21 and MARCXML. For MODS and HTML, a template was created for each metadata format and the application adds the supplied information to the corresponding element in the template, according to the mapping. Mappings from the elements provided in the application to MODS and Schema.org semantics were relatively easy due to the application's simple set of elements, each with clearly defined meanings.

### Technical Considerations

The application was created in JavaScript, which can handle enough complexity to fulfill the library's needs, and is easy to modify for additional functionality. The JavaScript code runs on the user's computer (the user's web browser must have JavaScript enabled to use the application) and is computationally simple enough that it can run quickly on any computer. The application works best in Google Chrome.

Currently the web application is hosted on one of the web servers in the UIUC Library, which can be used by anyone who knows the URL.<sup>26</sup> All source files and a simple instruction document are available in GitHub with an MIT

license, a standard license for open source software, so any institution can use and modify the code for their needs.<sup>27</sup> For institutions that do not have programmer support, Metadata Maker now has an entrance page.<sup>28</sup> By adding institution specific information on the entrance page, Metadata Maker populates this information in the form and adds it into the output metadata as a default value. Modifying the source code is simple enough that after the initial development of Metadata Maker, CAM was able to customize the application for cataloging backlogs of theses and dissertations, data sets, and government documents.<sup>29</sup> Another department in the UIUC Library that is cataloging backlogs of serials has been modifying the code to allow the application to work with serial materials.

### Testing the Application

After version 1.0 of the application was stable, the Foreign Language Cataloging Specialist conducted user testing with six CAM staff members. The user testing was focused on the MARC21 format metadata creation since it is the most used metadata format in the UIUC Library. Testers included two student workers, whose daily work is physical processing of materials, two graduate assistants with copy cataloging experience, and two hourly staff who have experience in copy cataloging and catalog maintenance work. Testers were given a variety of monographs in Western European languages, including English, which required original cataloging. With little instruction, they were asked to complete the form for each item and create metadata in all formats. Staff were asked to follow a standard naming convention for their records, and they wrote the file name on a streamer placed in the corresponding book. The files created by each user were collated in a shared network drive for the project. More than 240 records were created as part of this initial test; five of the users created between 5–8 records per hour, while one user created more than 10 records per hour.

After records were created, the Foreign Language Cataloging Specialist evaluated a sample of 88 records created in MARC21 format. It was decided at this point that only MARC21 records would be evaluated because of the format's complexity. The Foreign Language Cataloging Specialist imported all MARC21 records created by testers to a cataloger's local save file in OCLC Connexion and reviewed each record against the corresponding item. To preserve the UTF-8 symbols (for example, a copyright symbol) and non-Roman characters, the import record character set was selected as UTF-8 in OCLC Connexion. For all records, the Foreign Language Cataloging Specialist corrected any incorrect information, fixed errors in both coding and transcription, created subject headings based on the provided keywords, and controlled name headings if authorities were available in the LC Name Authority File

(LCNAF). All records were assigned appropriate classification based on the item, for example, Dewey Decimal, Library of Congress, or a local classification designed for special collections.

Because of the variety of resources used for the test, some materials needed more enhancements to create a complete record than others. On a case-by-case basis, the Foreign Language Cataloging Specialist decided if a record required additional information and upgrading to a full-level record. Upgrading of records was limited primarily to newer acquisitions and included the addition of various fields including statements of responsibility (subfield \$c of the data field 245), information about bibliographies and indexes (data fields 504 and 500), language notes (data field 546), etc. After the records were edited, they were added to OCLC's WorldCat database and exported to the library's ILS for normal cataloging workflow.

#### Identified System and User Errors

The Foreign Language Cataloging Specialist maintained a spreadsheet to track the changes made to each record reviewed and noted if the change was made to fix a system or user error. If the change was because of user error, notes about possible training to address the error were included. After review of the records, the list of errors by MARC field were collated in the spreadsheet and divided by error type (e.g., system or user). The system errors can be found in table 3 and the user errors in table 4.

All system errors have been addressed within the application, including support for nonfiling indicators for English and French language materials. Support for all articles of foreign languages listed in appendix C of the RDA Toolkit has been discussed, but there are no plans for implementation soon.<sup>30</sup> Articles of foreign languages not included in the list will be added to the tool when they are identified. Coding of dates has been modified to supply a publication date in brackets in subfield \$c of data field 264 if only a copyright date is available, and the fixed fields 008 positions 06 and 07-14 are coded to include both dates. Diacritic support has been added with the use of a pop-up window that allows the user to select the diacritic to be added to the preceding character, similar to the method used in OCLC Connexion. If a user checks the "Unlisted" box for any publication information, the appropriate "[ . . . not identified]" phrase is now mapped to the corresponding subfield in data field 264, and an "unpaged" box has been added to the extent element that maps the phrase "1 volume (unpaged)" to subfield \$a of data field 300.

**Table 3.** System Errors Identified through Testing

MARC Field	System Error
008/06, 008/07-14	Lack of coding for both publication and copyright dates when dates are the same
Descriptive fields	Lack of diacritic support
100	Incorrect punctuation Relationship designator in ‡c instead of ‡e
245	No support for non-filing indicators
264_1 ‡a, b, c	Element not mapped if "unlisted" in form
264_1 ‡c	No implied publication date if only copyright date provided
300 ‡a	No support for unpaginated materials
700	Incorrect indicators

**Table 4.** User Errors Identified through Testing

MARC Field	User Error
008/35-37	Incorrect language selected
008/33	Incorrect literary form selected
008/15-17	Did not select state/province when appropriate
020	ISBN recorded with hyphens
100	Incorrect form of name
245	Title recorded from cover instead of title page Incorrect capitalization Typographical errors
250	Edition statement not recorded Edition statement translated into English
264_1 ‡a, b	Incorrect information recorded
264_1 ‡c	Date taken from CIP information Date implied from copyright date
300 ‡a	Lacked preliminary page sequences Incorrect information
300 ‡c	Incorrect measurements for oblong materials
700	Lack of additional names for editors and illustrators

The majority of user errors could be addressed by basic training. While the ideal is to keep training to a minimum, simple instructions focused on choosing the preferred source of information and how to record information would increase the quality of metadata created. New users should also be walked through the application to point out the options for literary form (fixed field 008 position 33), to note the difference between the "Place of publication" and "Country of publication" elements and the state/province exception in the latter, and to answer any preliminary questions the user may have. While each element in the application includes help text, providing users with some instruction up front ensures that they know the basic expectations and allows them to easily refer to the help text when a refresher

is needed. As with any project or workflow of this type, time spent training and providing context at the outset produces better results and a more engaged user.

A surprising result of the test was a problem that arose when users were asked to create metadata for materials in languages they did not know. While many catalogers have basic bibliographic knowledge of numerous Western European languages, it was incorrect to presume this same level of knowledge in users who will use Metadata Maker. Lack of language expertise resulted in errors with forms of names, transcription (based on the idea that one with language knowledge could more easily catch typographical errors), lack of or incorrect edition and publication information, and keywords that were not useful for the assignment of subject headings. The prevalence of such errors identified during the test phase underscores the need for libraries to select users with language expertise appropriate to the materials to be cataloged with the application.

#### “Note to cataloger” Field

Along with tracking changes and errors, the Foreign Language Cataloging Specialist also kept a record of the information added to the *Note to cataloger* field, which was used during the test in a variety of ways. When testers were asked to handle materials in unfamiliar languages, they often entered data indicating that they were unsure about a certain piece of information. The element was also used to list additional names appearing on the item and their roles, which some testers did not think should be entered into the *Author* element. While it was expected that additional contributors would be entered under *Author*, which includes the option to add multiple values for *Author*, this use emphasized the need to reconsider the element name and the addition of dropdown boxes to allow the user to select the appropriate role.

In some cases the *Note to cataloger* element was used to enter information that added to the overall quality of the metadata but was not addressed elsewhere in the form, including information about related works, translations, the presence of multiple languages, and multiple places of publication. Because the data in this element is currently mapped to the MARC21 data field 500 (General Note), an institution or library could decide to use the element more systematically and give users instruction on the type of notes to be entered. This could serve to add to the quality of metadata without adding additional elements and mappings to the different metadata formats. Depending on the scope of a given project, these notes could be used by the cataloger to upgrade records, or they could be left in the data field 500 as additional information for both patrons and other catalogers.

## Conclusion

Metadata Maker was developed for anyone to create metadata for the discovery and access of hidden collections or backlogs that currently lack descriptive metadata. Initial testing has revealed that the application is easy to use and creates metadata that supports resource discovery, access, and management. The application also supports diacritics and Unicode non-Roman language encoding that would greatly help foreign language cataloging workflows. If the item in hand has the appropriate information, users could potentially create full-level metadata using Metadata Maker in four different formats.

As more metadata and cataloging authority sources make their resources available through application programming interfaces (APIs), the Metadata Maker project team is also working to improve the application by identifying available services and exploring ways to add them in new versions. As a first step, Metadata Maker version 1.1 includes suggested FAST headings as an option for the *Keywords* element, which allow users to add more controlled subject terms into the metadata, greatly improving discoverability through subject-based faceted browse services. The addition of the Virtual International Authority Files (VIAF) service is also being considered for all names included in the application, so that users can add authorized forms of names directly from the Metadata Maker web form. Both of these improvements also allow for future linked data capabilities. Additionally, the WorldCat Metadata API can be added for direct import of metadata into WorldCat, if the UIUC Library decides to take that route in the future. Because the application was created as an open source project, any institution can modify the codes as needed. The project team hopes that Metadata Maker will grow into a tool that is used and improved upon by the library community, and that it will benefit the community as well as library users.

## References

1. Jeanne M. K. Boydston, and Joan M. Leysen, “ARL Cataloger Librarian Roles and Responsibilities Now and In the Future,” *Cataloging & Classification Quarterly* 52, no. 2 (2014): 229–50; Matthew P. Long, and Roger C. Shonfield, “Ithaka S+R US Library Survey 2013,” accessed May 19, 2015, [www.sr.ithaka.org/wp-content/uploads/2015/08/SR\\_LibraryReport\\_20140310\\_0.pdf](http://www.sr.ithaka.org/wp-content/uploads/2015/08/SR_LibraryReport_20140310_0.pdf).
2. Myung-Ja Han et al., “Final Project Report for Making Metadata Maker,” accessed March 1, 2015, [www.library.illinois.edu/committee/exec/innovation\\_fund/innovation\\_fund\\_proposals/2014-2015/Innovation\\_Grant\\_Metadata\\_Maker\\_Final\\_Report.html](http://www.library.illinois.edu/committee/exec/innovation_fund/innovation_fund_proposals/2014-2015/Innovation_Grant_Metadata_Maker_Final_Report.html). See also, <http://quest.library.illinois.edu/marcmaker>; OCLC is in development of the *Low*

- Barrier Metadata Creation* tool that also supports minimal-level metadata creation. The tool is being tested by several academic libraries and vendors as of June 2, 2015.
3. Boydston and Leysen, "ARL Cataloger Librarian Roles and Responsibilities Now and In the Future," 229–50.
  4. *Ibid.*, 235.
  5. Long and Shonfield, *Ithaka S+R US Library Survey 2013*, 29–31.
  6. *Ibid.*, 30.
  7. *Ibid.*
  8. Richard Sapon-White, "Subject Analysis Training for Cataloging Paraprofessionals: A Model for Ongoing Learning and Support," *Technical Services Quarterly*, 26 no. 3 (2009): 183–93.
  9. *Ibid.*, 186–88.
  10. Barbara M. Jones, "Hidden Collections, Scholarly Barriers: Creating Access to Unprocessed Special Collections Materials in North America's Research Libraries: A White Paper for the Association of Research Libraries Task Force on Special Collections," (white paper, Association of Research Libraries Task Force on Special Collections, 2003), [www.arl.org/storage/documents/publications/hidden-colls-white-paper-jun03.pdf](http://www.arl.org/storage/documents/publications/hidden-colls-white-paper-jun03.pdf). See also Barbara M. Jones, "Hidden Collections, Scholarly Barriers: Creating Access to Unprocessed Special Collections Materials in North America's Research Libraries," *RBM* 5, no. 2 (Fall 2004): 88–105.
  11. *Ibid.*, 4.
  12. Judith M. Panitch, "Special Collections in ARL Libraries: Results of the 1998 Survey Sponsored by the ARL Research Collections Committee," (Washington, DC: Association of Research Libraries, 2001), accessed June 2, 2015, [www.arl.org/storage/documents/publications/special-collections-arl-libraries.pdf](http://www.arl.org/storage/documents/publications/special-collections-arl-libraries.pdf); Jones, "Hidden Collections, Scholarly Barriers," 4.
  13. Karen Snow, "A Study of the Perception of Cataloging Quality among Catalogers in Academic Libraries" (PhD diss., University of North Texas, 2011).
  14. *Ibid.*, 74–75.
  15. *Ibid.*, 98, 129, 160.
  16. *Ibid.*, 114.
  17. Roy Tennant, "MARC Usage in WorldCat: Zoomable Starburst of all MARC Tags—By Formats: Books," accessed October 1, 2015, <http://experimental.worldcat.org/marcusage/viz/starburst.html>.
  18. Karen Calhoun et al., *Online Catalogs: What Users and Librarians Want* (Dublin, OH: OCLC, 2009), accessed May 29, 2015, [www.oclc.org/reports/onlinecatalogs.en.html](http://www.oclc.org/reports/onlinecatalogs.en.html).
  19. University of Illinois at Urbana-Champaign, "Emblematica Online: Resources for Emblem Studies," accessed October 24, 2014, <http://emblematica.grainger.illinois.edu>.
  20. Timothy W. Cole, Myung-Ja K. Han, and Jordan Vannoy, "Descriptive Metadata, Iconclass, and Digitized Emblem Literature," *Proceedings of the 12th Annual Joint Conference on Digital Libraries* (Washington, DC: ACM, 2012): 111–20.
  21. "SPINE Metadata Schema," XMLSPY, last updated October 2, 2011, accessed May 29, 2015, <http://diglib.hab.de/rules/schema/emblem/emblem-1-2.xsd>.
  22. Library of Congress, Network Development and MARC Standards Office, "Appendix C—Minimal Level Record Examples," accessed October 24, 2014, [www.loc.gov/marc/bibliographic/bdapndxc.html#book](http://www.loc.gov/marc/bibliographic/bdapndxc.html#book); Snow, "A Study of the Perception of Cataloging Quality among Catalogers in Academic Libraries," 114.
  23. Library of Congress Network Development and MARC Standards Office, "MARC Code List for Languages," last updated September 8, 2011, [www.loc.gov/marc/languages](http://www.loc.gov/marc/languages); Library of Congress, Network Development and MARC Standards Office, "MARC Code List for Countries," last updated April 4, 2008, [www.loc.gov/marc/countries/](http://www.loc.gov/marc/countries/).
  24. Myung-Ja K. Han, "How to Use Marc Maker," last modified April 9, 2015, <https://github.com/dkudeki/metadata-maker/blob/version-1.1/How%20to%20use%20Metadata%20Maker.docx>.
  25. "University of Illinois Library Catalog Datasets," University of Illinois at Urbana-Champaign Library, last updated October 10, 2014, <http://catalogdata.library.illinois.edu>; OCLC, "OCLC Adds Linked Data to WorldCat.org," press release, June 20, 2012, [www.oclc.org/news/releases/2012/201238.en.html](http://www.oclc.org/news/releases/2012/201238.en.html).
  26. "Metadata Maker," University of Illinois at Urbana-Champaign Library, last updated on August 27, 2015, <http://quest.library.illinois.edu/marcmaker>.
  27. "MIT License," accessed June 4, 2015, <http://choosealicense.com/licenses/mit/>; GitHub Metadata-Maker, accessed June 4, 2015, <https://github.com/dkudeki/metadata-maker>.
  28. "Metadata Maker Entrance," University of Illinois at Urbana-Champaign Library, last updated on September 23, 2015, <http://quest.library.illinois.edu/marcmaker/entrance>.
  29. "Metadata Maker for Theses," University of Illinois at Urbana-Champaign Library, last updated on September 28, 2015, <http://quest.library.illinois.edu/marcmaker/theses>; "Metadata Maker for Data Set," University of Illinois at Urbana-Champaign Library, last updated on April 21, 2015, <http://quest.library.illinois.edu/marcmaker/dataset>; "Metadata Maker for Government Documents," University of Illinois at Urbana-Champaign Library, last updated on April 21, 2015, <http://quest.library.illinois.edu/marcmaker/govdocs>.
  30. "RDA Toolkit, Appendix C: Initial Articles," American Library Association, August 11, 2015, [http://access.rdatoolkit.org/rdaappc\\_rdac-27.html](http://access.rdatoolkit.org/rdaappc_rdac-27.html).