Notes on Operations

Descriptive Metadata for Digitization of Maps in Books: A British Library Project

By Kimberly C. Kowal and Christophe Martyn

Hidden special collections are increasingly being made visible and accessible by small digitization projects. In the project described in this paper, the British Library employed existing library standards and systems to accomplish key functions of a project to digitize a selection of maps contained within rare books. The integrated library system, using the Anglo-American Cataloguing Rules (AACR) and Machine-Readable Cataloging (MARC) format, acted as a lynchpin, linking directly bibliographic descriptions of both the original and the digital copies of the map, the book containing the map, the digital image, and preservation data and strategy, making the items widely searchable and visible while uniting them with the broader collections.

In tandem with the surge of mass book digitization projects has been a movement to highlight small special collections with digitization and cataloging. The Library of Congress (LC) Working Group on the Future of Bibliographic Control recommended as a priority enhancing access to rare, unique, and special hidden materials, encouraging digitization and creation of detailed descriptions, as well as integrating access to these materials with wider institutional holdings.[1] With the capabilities of today's library systems, a surprisingly large number of these tasks are possible in many libraries using existing library skills and resources.

In the project described in this paper, the British Library (BL) employed existing and emerging library standards and systems to accomplish key functions in a project to digitize a selection of maps and views contained within rare books. While the project involved a number of stages and areas of expertise, this paper will explicate the manner in which the authors handled the need for descriptive metadata identifying the item and its source, documenting copy-specific attributes, and making the record and its digital surrogate accessible. The main library system in the BL, the Aleph 500 integrated library system (ILS) produced by Ex Libris, acted as a lynchpin, linking directly bibliographic descriptions of both the original and the digital copies of the map and the book in which the map appears, the digital image files captured, and the preservation strategy, making them widely searchable and visible while uniting them with the broader collections. This project represents the first use in the BL of the digital asset module in Aleph.

Description of the BL Project

The Vulnerable Collection Items Project was undertaken at the BL to select, digitize, and collect metadata for maps held within the rare printed books collection. Following thefts of valuable maps contained within books from multiple institutions that included the BL, it was thought that a method should be developed

Kimberly C. Kowal (kimberly.kowal@bl.uk) is Lead Curator, Map Library, The British Library, London; **Christophe Martyn** (chris.martyn@bl.uk) is System Projects Coordinator, Systems Management, The British Library, Boston Spa, United Kingdom.

to firmly identify the unique copies of rare and important BL holdings to better protect valuable collection materials considered vulnerable. The resulting process combined sets of high-resolution security photographs, bibliographic metadata to describe the physical object (which includes copy-specific descriptive metadata such as condition descriptions), metadata for the digitized image, and linking this and the image to the bibliographic metadata. This enabled the highest possible level of identification of distinguishing features that existing BL systems can accommodate, improving the security of the selected maps.[2]

The original, security-oriented project aims eventually blossomed into something of more universal use and wider research value. Having acquired digital photographs of the collection items and associated metadata, it became clear that sharing the information would contribute to accomplishing other BL strategic priorities. The project could serve to answer the library's security concerns while enhancing user access to the collections by providing publicly accessible metadata for, and images of, the maps under consideration. The advantages of revealing these hidden collections were deemed to far exceed the potential pitfalls inherent in extending the project's aims.

Literature Review

In the plethora of funded digitization projects throughout educational and cultural sectors, visual collections identified as "otherwise hidden" have been well represented in recent years. Methods for capturing metadata during digitization projects for such special collections have been plentiful in the current literature, representing manuscripts, ephemera, fanzines, remotely sensed imagery, original art, architectural images, posters, and postcards.[3]

Maps are no exception to this attention, with numerous scanning products using a variety of standards and methods for metadata capture evident on the Web. The American Library Association Map and Geography Round Table Map Scanning Registry, ongoing since 2006, is the primary online listing of map scanning projects.[4] This constantly updated source provides outline information (prepared by the project owners) about the projects, describing the content, technical standards used, and metadata captured. Though most of the projects represented are either not collecting metadata or have not provided this information in the registry, those that have done so list the Federal Geographic Data Committee Metadata Standard (FDGC), Machine-Readable Cataloging (MARC) 21, or Dublin Core (DC) standards, which are widely adopted by metadata librarians for digitizing projects and born-digital data collections.

MARC as a Metadata Solution

The use of MARC and the Anglo-American Cataloguing Rules (AACR) has been a less popular approach for capturing descriptive metadata for special format digitization projects. MARC as a tool in digitization projects has been criticized in the past for being "too complex, requiring highly trained staff and specialized input systems," and for being too focused on print material and not extendable for digital collections.[5] More recent reviews and comparisons have looked upon MARC more favorably. Beall outlines twelve criteria for comparing metadata schema, suggesting numerous advantages for the use of MARC in library projects.[6] Significant among Beall's criteria is the availability of systems and software to support any given metadata scheme, meaning that MARC metadata can be created and searched in library ILSs, a desirable feature often taken for granted. Layne

reported as early as 1991 the use of the MARC format for a digitization project of images in medieval manuscripts.[7] Recognizing that MARC is very often not the method of choice for manuscript materials, her primary question related to the usefulness of MARC for this purpose, which was to provide description and access, while applying widely known and accepted standards. She concluded that the flexibility of the format was effective, and it continues to be used. Other small, special format digitization projects described in the literature using MARC include the joint University of Pennsylvania Library–Cambridge University Library project in 2006 to create online catalog records and an image database on the Web of dispersed manuscript fragments and the mixed collections of ephemera (Pennsylvania German broadsides and Fraktur) taking place at Pennsylvania State University Libraries.[8] The former selected MARC principally to integrate the descriptive metadata into the existing library system while the website uses a crosswalk to convert the records to DC. The latter focussed on the challenges of using the multiple sets of cataloging rules accompanying formats of monographic broadsides, graphic materials, and manuscripts.

These cases represent comparatively small projects limited to finite collections. MARC is also applied for ongoing, nonproject-based digitization. Two major institutions using MARC for ongoing maps digitization are the LC Geography and Maps Division, as part of the American Memory Project (http://memory.loc.gov/ammem), and the Harvard Map Collection (http://hcl.harvard.edu/libraries/maps). Both of these use MARC records for descriptive metadata and to provide access to scanned map images from the online public access catalog (OPAC) through hyperlinking. When the link in the OPAC is selected, an external viewer is launched that allows

interactive features with the image, such as panning and zooming.

Previously published research relating experience with the use of MARC for special format digitization projects enriches understanding of the benefits of using MARC as a format and the challenges and methods of interpreting *Anglo-American Cataloguing Rules*, 2nd ed. *(AACR2)* cataloging standards for such collections.[9] Thus far, however, no detailed practical reports on the way in which MARC analytics can be applied to reveal visual materials hidden within another bibliographical unit (e.g., maps in books) or how bibliographical data would be structured to accommodate this have been published.

Other Metadata Solutions for Digitized Maps

Many digitization projects are employing metadata standards specifically designed to capture information about digital image data, and these schemas reflect the flexibility possible in the new and continually emerging systems used to manage them. Projects handling cartographic materials, in common with wider practice, use any number and combination of standards and methods of capturing metadata, including the well-established DC, MARC, Federal Geographic Data Committee (FGDC), Encoded Archival Description (EAD), Metadata Encoding Transmission Standard (METS), and Metadata Object Description Schema (MODS). The projects including maps are too numerous to detail comprehensively, so a select few are highlighted to illustrate the diversity and flexibility of solutions being developed.

In many cases, a defined scheme is adapted to the project. In the case of the Collaborative Digitization Program (originally the Colorado Digital Program), no one standard was deemed appropriate; the participating institutions' existent metadata schema were all mapped to a minimal DC element set in order to facilitate efficient crosswalking of metadata from the archival, museum, and library collaborators.[10] The same type of amalgamation was applied by the project librarian Nicholas Graham for North Carolina Maps (www.lib .unc.edu/dc/ncmaps), a collaborative project merging images and records from library and archives catalogs in both MARC and EAD.[11] The data from existing catalog records were downloaded to a spreadsheet, additional fields were added, and the plan is to eventually export these data to MODS. Such a method requires consistent mapping between the various metadata fields, and Graham's work crosswalking between four standards is invaluable.

In other cases, more than one set of metadata is captured, allowing different standards for different purposes. METS, an Extensible Markup Language (XML)–based schema for packaging related sets of digital objects, was used for digitised Sanborn maps at University of Colorado at Boulder Libraries, but only after MARC records were created for the digital and analog versions in the library catalogs, with the data then converted to XML.[12] In combination with locally developed tools, the project used MarcEdit, a freely available utility developed by Terry Reese at Oregon State University for batch editing and converting MARC between formats.[13] The same tool was used by Brenner in her innovative project with the Oregon Sustainable Community Digital Library (http://oscdl.research .pdx.edu) to merge metadata from disparate contributors, display scanned materials through Google Earth, and provide MARC metadata, all directly from the library's OPAC.[14]

Scanned maps that are converted to geospatial data, as in McGlamery's monumental distributed project of scanned and geo-referenced topographic maps of the Austro-Hungarian Empire, require more specialised content standards.[15] Although the International Organization for Standardization standard for geographic information (ISO 19115) was considered, FGDC's Content Standard for Digital Geospatial Metadata, with its antecedents in MARC, was ultimately selected, and a customised application was developed for metadata input.

These standards are used effectively with a host of new commercial software products devised to manage metadata collection, discovery, distribution, and display of digitized images. Referred to collectively as Digital Visual Information Management, these systems include library OPACs, content management systems, digital asset management systems, and digital repositories.[16]

Managing Digitized Content in the BL

A number of approaches are currently being taken to manage digitized content in the BL. Although MARC and *AACR2* are used for the majority of BL cataloging, the BL has used different standards for specific circumstances, and a combination of in-house BL work and components provided by third parties for metadata and systems is usually used.

The BL Application Profile (BLAP), an extended DC–based declaration of descriptive metadata terms encoded as XML, was developed to support a high-level cross-searching facility among BL resources of different types and with different metadata formats, and has been used in several large BL digitization projects. An early implementation was Collect Britain (www.collectbritain .co.uk), one of the largest digitization projects thus far carried out by the BL, the aim of which was to digitize a selection of historic content from several BL collection areas. BLAP metadata was stored in an Structured Query Language (SQL) Server 2000 database and digital objects in a file

store, using content management systems to upload and deliver Web content. Another large project, British Newspapers 1800–1900 (www.bl.uk/reshelp/findhelprestype/news/newspdigproj/ndproject), digitizing up to two million pages of British national, regional, and local newspapers, used a customization of BLAP and an Open Archives Initiative for Metadata Harvesting (OAI-PMH) data provider service for interoperability requirements within third party content management systems. BLAP also has been used to describe sound recordings in the Archival Sound Recordings project (http://sounds.bl.uk) and to provide a metadata standard for use with BL Web resources.

Many small, discrete digitization projects fall into the Themed Collections Programme, defined as "systems developed using a standard software architecture designed to hold varied and complex data and allow it to be searched, edited, and presented in various 'themed' ways, usually on the internet."[17] These were devised to provide cataloging and a search interface for collections considered incompatible with the ILS; bibliographic data is stored in XML in SQL Server 2005. The Themed Collections has been used successfully for several BL projects that mix metadata, text, and images. These include databases of, for example, Renaissance Festival Books (www.bl.uk/treasures/festivalbooks/homepage.html), Database of Italian Academies (www.bl.uk/catalogues/ItalianAcademies), and Historic Photographs (www.bl.uk/onlinegallery/features/photographicproject/index.html). In addition to resource discovery through the BL website, the system facilitates data exchange with other organizations, links to digital objects, and item requesting. METS has already been used in the BL as a "wrapper" for the Archival Sound Recordings project (http://sounds.bl.uk) and is now also being used to package the various types of metadata associated with e-journals (e.g., MODS, PREservation Metadata: Implementation Strategies (PREMIS)). The use of METS will no doubt be extended to other content types.

OAI-PMH is currently being investigated by the BL to harvest data from digital objects stored in the BL's Digital Library System so that it can be used for a variety of resource discovery initiatives such as the European Union–funded Europeana project (http://dev.europeana.eu). OAI-PMH is still being used in the BL only for specific projects such as these, and although not ready to be used for this project, the BL plans to expand its usage into more general areas.

Selection of a System

The need to capture and organize descriptive metadata to accompany the digitized map images meant that, in order to be effective, the system needed to

- ingest the description of the map, its bibliographic source, and the individual copy condition;
- accommodate electronic searching and access to the records and potentially images, ideally linking the two; and
- ensure institutional long-term maintenance, preservation, and technical support.

At first glance, the Themed Collections system seemed appropriate, since it had been used at the BL in the past to manage images and metadata associated with special collections. With further examination, however, the ILS was chosen, for several reasons:

- Avoiding the unnecessary creation of a new software or website-specific database for this project was considered of paramount importance. The use of Themed Collections software and hardware would have necessitated building and populating a new database, which would have been costly and time-consuming.
- Bibliographic records for the books containing the images were already in the ILS, as were some records representing individual maps contained within books or atlases.
- The ILS supported the use of analytical bibliographic records to describe discrete elements contained within bibliographic units (e.g., a map within a book) by means of "child" analytical bibliographic records linked to the "parent" bibliographic record representing the work containing the map.
- The ILS possessed functionality to link the digital images to the metadata, so it could in theory present a complete representation of the image to the user.
- Resource discovery by the public was already possible through the OPAC and allowed the flexibility to make the record public (or not). This was an important consideration as the project developed and the desired outcomes changed.
- Because the ILS is the core cataloging and resource discovery system used by the BL, future support for records created for this project was guaranteed.
- An infrastructure (the BL's Digital Library Programme) was already in place to support ingesting and preserving digital objects.

Selection of Standards

Several of the various standards described above appeared to be potential solutions, including MARC 21/

AACR2, BLAP, MODS/METS. The reasons why MARC 21/AACR2 was chosen included the following:

- DC-derived standards such as BLAP likely would have to be customized for this project, which would be time consuming, whereas the most commonly used standards in the BL, MARC and AACR2, could be employed without modification.

- AACR2 is the international content standard used by the BL to describe much of its collection, and it is the standard used in current cataloging. In addition to printed books, AACR2 fully supports cartographic materials.

- The MARC 21 bibliographic, authority, and holdings formats (all used by the BL) provide a way to express catalog records created according to AACR2 standards in MARC format, and provides some additional information (e.g., coded data and content designation) used by computers to enhance access.

- MARC is a proven standard; its stability and its granularity for descriptive metadata recommended it. With these qualities it can both operate well in current systems and easily be migrated in the future.

- MARC is continuously growing to accommodate new technological advancements, and so is equipped to handle the necessary hybrid of print and digital information.

- MARC contains data elements for recording preservation actions.

- MARC supports the expression of relationships between related items.

- MARC and AACR2 enable the recording of information specific to particular copies of a work. This means that unique

characteristics of the map could be recorded, providing obvious benefits for collection security.

- MARC and AACR2 are at the heart of mainstream BL cataloging. Thus records created or reused for this project using those standards would follow the same development path as most other BL catalog records (e.g., forthcoming moves from AACR2 to its successor, Resource Description and Access (RDA), and from MARC 21 to XML–based MARC formats).

In addition to AACR2, its collateral publication, *Cartographic Materials: A Manual of Interpretation for AACR2*, was used as the primary authority consulted for reference, along with "MARC 21 Format for Bibliographic Data."[18] Other standards used included *Library of Congress Subject Headings* and the NACO (Name Authorities Cooperative) Authority File.[19] Although AACR2 was used to construct the bibliographic records representing the images, the host book records would most often not be constructed according to AACR2 because they were created long before AACR2 was introduced.

Because entire books were not scanned but only the maps, structural metadata was simple and could be noted in MARC. The level of technical metadata, included in the Aleph Digital Asset Management (ADAM) metadata record representing the raw images, was deemed sufficient.

Project Implementation

Although the use of the ILS as well as MARC and AACR2 appeared to be the most suitable approach, several aspects were untested. For example, the BL had until then cataloged below the level of the item only in the case

of conference proceedings and did not yet have a policy for recording copy-specific information. Additionally, the ADAM module, a priced add-on option to Aleph available beginning with version 16.03 that operates within the cataloging module, had been acquired by the BL but not exploited extensively; it allows small-scale (i.e., not a digital archive) management of digital objects within the Aleph environment. This project therefore presented a unique opportunity to test the feasibility of applying the BL's existing systems and standards to manage these complex facets of the project as well as a challenge to adjust local policy, technology, and practice to accomplish these ends.

The authors wished to use existing resources in the library, in terms of established bibliographic standards and technology, to integrate materials with the BL's larger holdings and to increase the items discoverability to users, whether they are searching for a citation or the digital image itself. The metadata structure, format, values, and content needed to fit into the established standards of the larger institution to ensure that it would be supported in future potential changes, such as migrations in the library system, shifting library standards, Web access, and technology. Cataloging this unusual medium (early cartographic images contained within books and their digital manifestations) required expanding how the BL currently used the standards and system. For the metadata segment of the project alone, it was necessary to draw on support and advice from numerous units of the library, including British Collections and the Map Library, for staffing, curatorial insight, and project management and to ensure the most up-to-date practices were used for map cataloging and digitization; Systems Management for problem solving and technical support to enable the ILS to suit the

project's needs; Bibliographic and Metadata Standards and Data Quality and Authority Control to review and approve the template elements and functionality; and Resource Discovery and Applications Development for creative development of the system and policy decisions regarding access to additional modules.

Operational Phase One

During the first phase of the project, constituting an operational period of approximately nine months, more than three thousand maps of the world and of the Americas, produced by Europeans between the late fifteenth century and 1700, were selected for inclusion in the project. Figure 1 shows a typical map included as it sits within its containing volume. It portrays the mid-Atlantic coast, a region of intense interest to Western Europe in the seventeenth century, and is contained within a 1651 English text, the *Discovery of New Brittaine* (London: I. Stephenson, 1651), describing the "discovery of New Brittaine." The book will be included in the catalogs of most libraries that own it; the map (illustrated in figure 1) will not. In some cases, the maps selected already had skeleton catalog records in the system, whereas in other cases there was no catalog representation because the BL only selectively catalogs special format material (e.g., illustrations and maps) contained within books. The books in which the maps were held were already represented in the BL catalog, the majority with minimal records retrospectively converted from the printed catalogs.

Staff for the cataloging portion of the project initially consisted of three individuals. The map curator designed the templates for the records and coordinated with other relevant teams in the library for policy decisions, advice, and to arrange required
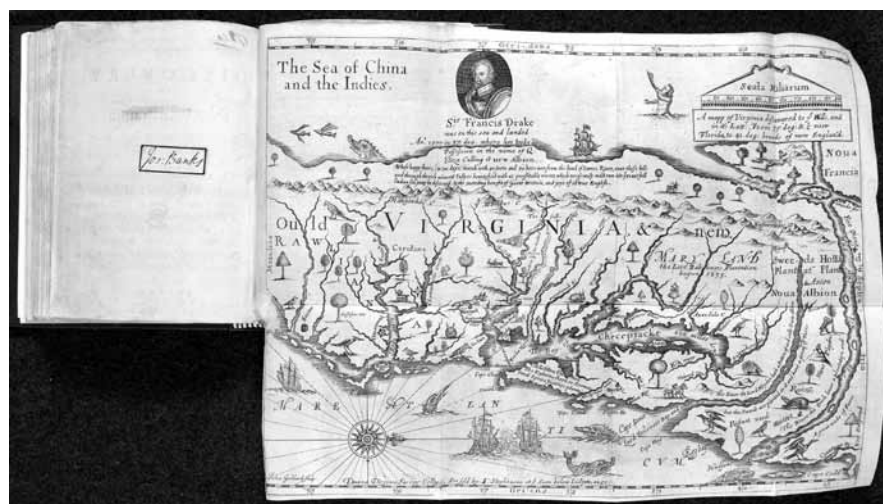
Figure 1. This English book, *The Discovery of New Brittaine,* contains a folded map within titled "a mappe of Virgina discovered to ye hills."

functionality. Two full-time project curators were employed to devote the majority of their time to cataloging. Between these two, exceptional expertise was brought on different fronts. One offered knowledge of the map literature and extensive experience with antiquarian maps, background appropriate to provide bibliographical reference citation notes, detailed condition descriptions and copy-specific information that would aid in identifying particular distinguishing features for each map. The other, a trained cataloger, brought adeptness with the library system, current standards, and the technology. Both had multilingual abilities, beneficial for handling the multitude of Western European languages. Above all, flexibility was an essential attribute for the project because the processes and technologies were in many cases new to the BL and had to be developed and coordinated as needs arose.
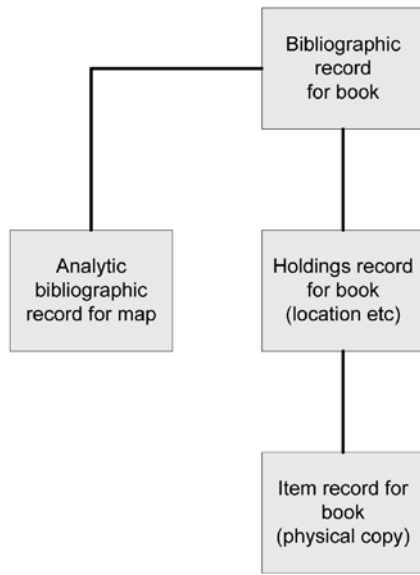
Descriptive Metadata: Map

The data structure for a book in the ILS takes the form of a MARC 21 bibliographic record for the book, a separate MARC 21 holdings record linked to the book record and containing data about the book's location as well as its physical condition and other aspects specific to the individual copy, and an Aleph-specific item record representing the physical copy. The item record is linked to the holdings record. The presence of an item record introduces some conceptual problems because the holdings record represents the individual copy in MARC; however, item records are essential because Aleph administrative functions are carried out against them.

The analytic bibliographic record representing the map does not have its own holdings record because holdings policy in the BL dictates that holdings below the item level may not be expressed. Instead, it is linked to the bibliographic record for the host item. This link enables viewing the location of the host item and requesting the item through the map analytic record, even though the holdings record is not linked directly to it. Figure 2 presents a model of this data structure.

The project team realized quickly that individually cataloging each map would be the most efficient means of identifying each and recording its location and context. Creating a new record
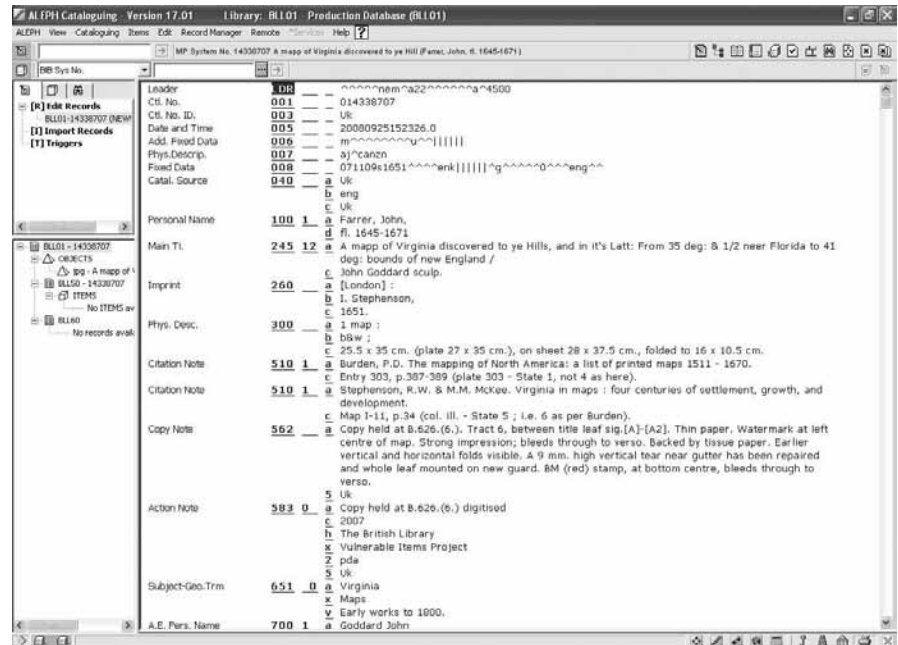
Figure 2. Model of the Data Structure Used

Figure 3. A Bibliographic Record for Map in Figure 1, As Seen within the Staff Interface of the BL's ILS

for the map using the MARC format and the *AACR2* standards for cartographic materials would capture the bibliographical information by which items might be searched. Additionally, the linking functionality offered by MARC and the ILS would properly express the relationship between the map and its host item.

The project template consisted of a set of core data elements. Common to other online catalog records for cartographic materials, they were already accommodated in the ILS and follow *AACR2* rules. Other fields, such as notes, added entries, and additional subject headings, were added when appropriate. The LC access-level record standard was not specifically considered for this project, but many data elements it includes were replicated.[20] Some data elements it contains were inappropriate for this project, for example the MARC linking fields 580 and 780 for reasons given in the following section.

The appendix presents a list of the fields in the template for the analytic bibliographic record for the map, with standard options and anomalies or

features specific to the project noted. Numerous other fields appear in the records. The Aleph Linker (LKR) field expresses the link between the analytic and the host record in the ILS. Among the copy-specific information fields used at the BL, the 562 Copy and Version Identification Note was employed, the development of which is described in the similarly titled section below. Additional required ILS–specific fields are also used. Figure 3 is an example of a completed record and describes the image shown in figure 1.

Linking the Analytic Record to the Host Record

Bibliographic records representing the pre–1700 books in which the maps were contained were already present in the ILS. This project did not require changing or upgrading these records in any way.

The analytic bibliographic records representing the maps had to be created where they did not already exist according to the standard described

above. Initially, the link between the child analytic map record and the parent book record was a tenuous one, built by manually entering the host and shelfmark as MARC fields 740 (added entries for related or analytical titles), 773 (information concerning the host item for the constituent unit described), and 852 (location) in each analytic record. The linking functionality afforded by the MARC linking fields, although technically possible in Aleph, was not used in the BL implementation; instead, the dedicated Aleph LKR field was used. The ILS can accommodate several different types of links between records using the LKR field; for example, it is used to link the holdings record to the bibliographic record. The link that was used for this project was the "Up/down" analytic link between bibliographic records of different levels, in this case between the parent book record and the child analytic map record.

Record system numbers, which are unique to each record in the ILS (and are used as the unique identifiers

throughout the project), are fundamental to the linking process. Although the LKR field has the functionality to create the links, the system number of the linked record in the LKR field identifies which record should be linked to which. The LKR field appears in the sample bibliographic record in figure 3.

Systems Management assisted in the development of a macro for inserting and populating an LKR field into the child record to generate a hyperlink between the two automatically, and this was integrated into the cataloging workflow. By entering the LKR field in one record, the ILS functionality generates reciprocal links between the parent and child records. One effect of this is to expand the location information given in the holdings record for the book into the analytic bibliographic record, facilitating requesting in the BL's onsite retrieval system.

Descriptive Metadata: Copy-Specific Information

Recording the condition description of each individual map was considered to be vital to identify unequivocally the unique copy of each image owned by the BL. This presented a challenge as basic copy-specific information was previously only recorded in item records and inconsistently and sporadically in bibliographic and holdings records. In response to the specific needs of this project and others throughout the BL, a new policy to enter copy-specific information at the holdings and bibliographic level in standard MARC fields was devised, allowing for such cases where analytics are used to represent part of a work.[21] Following this policy, the condition description is transcribed in the 562 field of the analytic bibliographic record instead of in the holdings record (as analytics may not have their own holdings records).

The content of the condition note (field 562) included the location of the map in the volume; description of paper, including location of watermarks or inequalities; printing, noting strength of impression, bleeding, offsetting, or plate marks; damage, such as stains, wormholes, tears, and repair work; or other markings including coloring or annotations. Because of the free-text nature of this field, a style sheet was developed with colleagues in the BL's Early Printed Collections to establish agreed vocabulary, abbreviations, and punctuation.

Identifying the copy as belonging to the BL within the relevant field was essential if the record was shared with another institution. Copy-specific fields in BL collection items are distinguished by the shelfmark of the copy being described preceding all content in the first subfield within each copy-specific field. This is because, although each note is linked to a single holdings record, it will be displayed in a bibliographic record that may be linked to several holdings. Also, as the details described only pertain to the BL copy, $5Uk is added to the end of each copy-specific field to identify the institution where the copy is held.

Including an indication that the map has been digitized (with the project affiliation) in each record was desirable to ease retrieval of maps included in the digitization project. This information was recorded in the area of Preservation and Digitization Actions (field 583), an area that could be compared to elements in preservation metadata schema. Use of this field is, as a matter of BL policy, normally reserved for use by conservation staff who use the dedicated Preservation and Conservation Management System, a separate instance of Aleph reserved for conservation work. Nevertheless, this field was used throughout this project so that the existence of a digital copy could be readily ascertained. This field is suppressed from public view by

specifying a particular indicator value, which the OPAC has been configured to take into account.

User Needs: Electronic Searching

Most of the relevant fields in the bibliographic record were already indexed and visible, and so searchable and viewable in the ILS configuration. As part of the work to compile the "British Library Policy for Copy-Specific Information," all relevant fields in the bibliographic and holdings record were reviewed, and relevant changes were made to the ILS to ensure that they were indexed and visible within the staff view.[22] This guaranteed searching and access to the records across the staff view. For display in the OPAC, fields considered sensitive or unnecessary for the public to view (e.g., 583) were suppressed from public display as described above. Interaction with the BL's requesting system was integrated in that, just as the parent book can be requested directly using the standard requesting function in the OPAC, the analytic may be requested in the same way. As previously stated, this is because the analytic also contains the location details of the parent book because of the functionality of the LKR field.

The needs of users searching the OPAC differed little from the original, internal-only audience. To staff and to the public educated in the structure and elements contained in library catalogs, searching by title, author, or subject are the expected retrieval methods for published materials. Within the OPAC, searches for records in the project may be limited by searching only within the Digital Items subset of the overall collection, or by searching for the Local Subject Heading, Scanned Maps, and Views. Outside of the OPAC, the new presence of these materials is highlighted within BL Help for Researchers webpages,

offering guidance on searching the ILS for maps, so it was not considered necessary to create additional webpages.

User Needs: Access to Images

The project team wanted to provide precise and quick retrieval of the images through a bibliographic search. The image and its metadata are linked to the bibliographic record. Immediate access to images serves several functions:

- It enables project staff to verify the ILS record and image authenticity as well as check the correct file naming assignment using a single system.
- It provides a visual finding aid (or a digital surrogate, depending on what is being investigated) to assist users in determining whether the material is of sufficient interest to warrant requesting the original volume.
- The storage of the image in the OPAC, with the bibliographic record and the requesting system, provides increased access in an immediate and familiar interface while sparing the fragile materials from unnecessary handling.
- The ADAM module allows capture of further technical metadata for rights management and access control with the image.

The ADAM module, which enables image files to be managed, delivered, and discovered, allowed "access images" (i.e., cropped, low-resolution JPEG images) to be attached to the analytic bibliographic records. Though the module is not yet widely used in the BL, permission was granted by the ILS Service Management Group to the project team for this initiative. A low-resolution access image of approximately 100–250 kb was created at the time of image capture for storage in ADAM. Like the TIFF
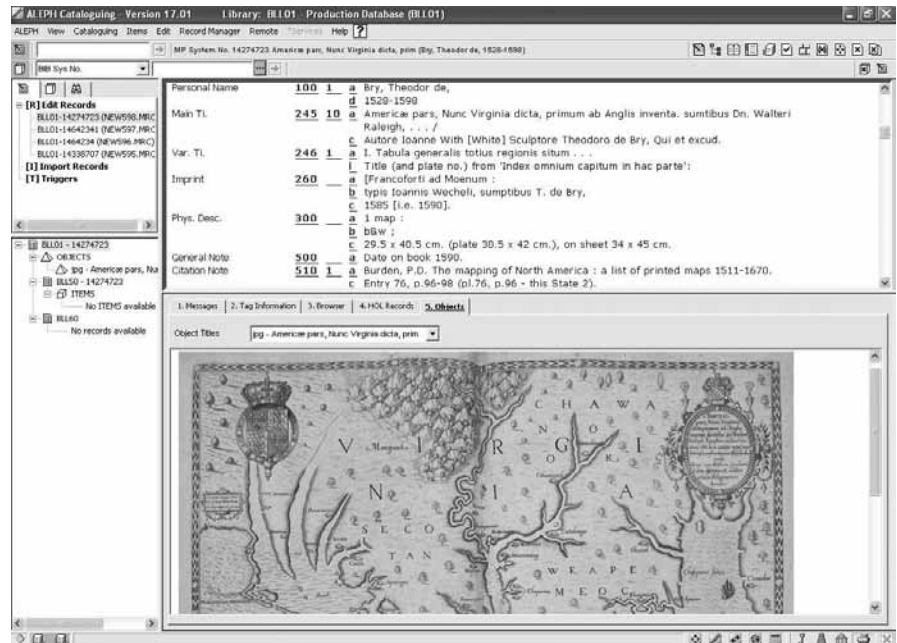

**Figure 4.** Screenshot of ILS Staff Interface, with the Digital Image Displayed alongside the Record

master file to be stored in the library's digital archive, it is named according to the maps' unique identifier, the ILS system number of the bibliographic record. These images are added manually to each of the records, thereby going through another process of ensuring the number, metadata, and record match. Figure 4 presents the object opened alongside the bibliographic record in the staff interface of the ILS.

In preparation for the records and attached images to be made visible in the OPAC, a batch service was run by the ILS team to create thumbnails for all of the objects. A screen shot of the appearance of records in the OPAC, with the thumbnail alongside, may be seen in figure 5. At the bottom of each record is an icon that links to the access-sized image of the file in a separate window.

The use of ADAM meant that some of the MARC fields that have become standard at other institutions for cataloging digital images and electronic reproductions were not used. In such cases, when an item is available

on the Web, the MARC 856 field (Electronic Location and Access for "information needed to locate and access an electronic resource") will contain the URL with an active hyperlink to the raster image. In the case of a reproduction of a print item, the record may include a second 007 field in the bibliographic record to describe the digital reproduction. Because the location of the file was not being provided, and the 006 field was applied to designate the secondary digital form, neither of these fields was used. The first template began with a MARC 530 (Additional Physical Form Available) note, but this was eschewed by the time of the second mutation because it repeated data already present in the record.

This tool is successful in managing the digital images for internal project purposes; for users, it provides an irreplaceable visual aid for an essentially graphic format that is difficult to visualize on the basis of the textual information supplied in a bibliographic record. Clicking on the thumbnail image produces a pop-up window
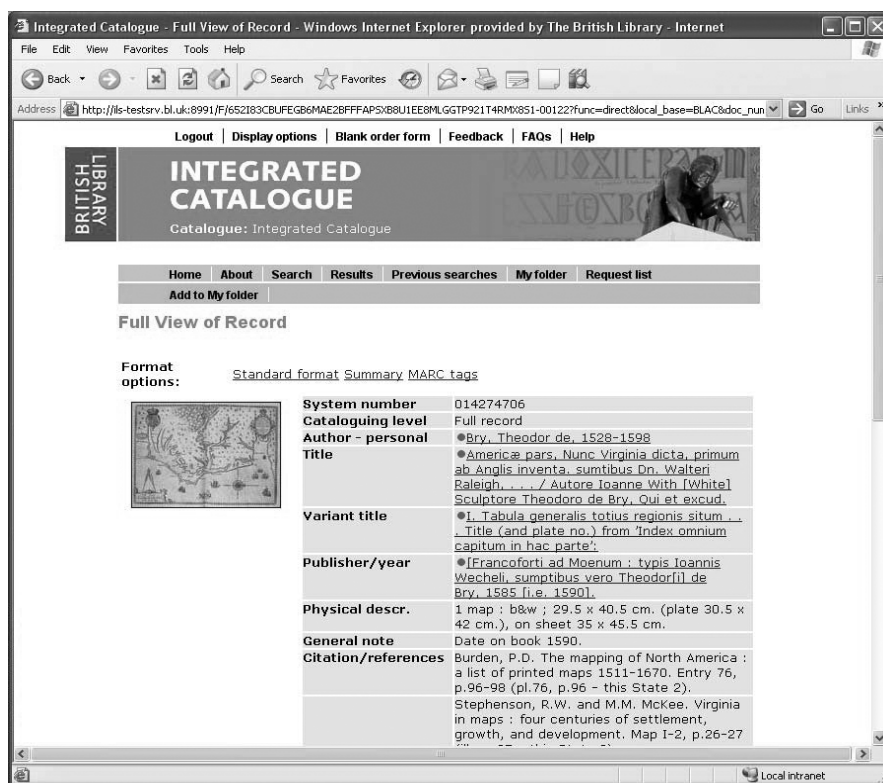
Figure 5. A User's View of the Record with a Thumbnail of the Attached Image in the OPAC (Cropped Image Represents a 1585 Map of Virginia within a 1590 Book)

with the enlarged access image, as in figure 6. Unlike most projects that make images available through a Web browser, however, there is no interactive functionality with the images. This could be a serious disadvantage if a user is interested in conducting research solely within the OPAC rather than using the images as a finding aid to consider if the original is relevant or worth consulting. The smallest typeface and other details on the map may not be sufficiently legible in the access images.

Along with item and holdings records, the digital object record created using ADAM forms part of the array of administrative data linked to the bibliographic records for the image and its host item. The ADAM record can contain metadata on copyright, access permissions to view the image, and technical details of the image. These are not as detailed or as sophisticated as they could be in

METS, but they served the purpose of this project.

The Effect of Changing Standards

Using established international cataloging rules and format standards throughout the library ensures flexibility with changing technology, leaving open the possibilities for alternate systems, expanded functionality, and secondary uses. Amid the rapid changes taking place in library systems, image management, and metadata standards, this is particularly important for small-scale project work, which can easily be left behind by larger changes. In the future, the BL will presumably move to an XML–based schema to represent descriptive metadata (though it is too early to speculate exactly what form this will take). The general trend in libraries is to move away from traditional OPACs and to replace these with Web interfaces that offer more

sophisticated and configurable display options and more user-controlled activities, such as tagging. An XML–based format is required for effective integration of bibliographic data in Web interfaces. Although MARC and the current data structures in the BL ILS are satisfactory, any move to an XML–based format will provide an opportunity to look at the data afresh with a view to improve its structure and display.

The BL is already planning to move to RDA, the successor to *AACR2*, in 2010. RDA is based on FRBR principles. In traditional cataloging, bibliographic units are described out of context; in FRBR, items must be described in context in a manner sufficient to relate the item to the other items making up the work. In the fullest implementation of RDA (Scenario 1), data is stored in a relational or object-oriented database structure that mirrors the FRBR and FRAD conceptual models. This more effectively supports what this project attempted to achieve—expressing the relationship between individual digitized images and their host work, documenting information about the image and making it accessible. The BL will initially implement Scenario 2 of RDA (currently scheduled for the first half of 2010), in which bibliographic and authority files are linked and which the BL ILS supports. This means that the introduction of RDA into the BL will not have any effect on this project.

Discussion

Several larger issues emerged during the time span of the project, which, rather than being discussed in detail here, will be noted as ongoing concerns.

Library Security

The issue of security arose as a result of the proposal to make the information available to the public.

Curators questioned whether exposure (i.e., making users aware of the existence of library materials) would make those items more vulnerable to theft. This attitude is an enduring one and counters widespread library practices such as cataloging, digitization, and creation of indexes, research guides, and finding aids. There has been extended discussion among professionals on library security of maps, a topic beyond the scope of this paper.[23] It is generally felt that a descriptive catalog record, especially one that includes copy-specific information, is a record of ownership and serves to protect materials.

The Role of the OPAC

Digitized images and associated metadata are often presented through separate, dedicated project websites, even when prepared by libraries with an OPAC. This may be because of limitations of library system technology or metadata structures in the past. Alternately, it could be attributed to the inevitable progression of research methods, user expectations, and information access in society, which can make OPACs, *AACR2*, and MARC 21 seem inadequate and obsolete.[24]

Responsibilities within the Institution

Though this paper only discusses a single element of the project (i.e., metadata capture), the project team was relatively successful in bringing together various departmental interests and input. This brought out questions, however, as to who should be doing what and for whom within the institution, given the number of tasks that were new and not neatly designated in job descriptions or by precedents. Frequently, libraries have dedicated staff for digitization, but even in the best cases the institutional infrastructure is relied upon to make it functional. This raises the question as to whether many librarians will move
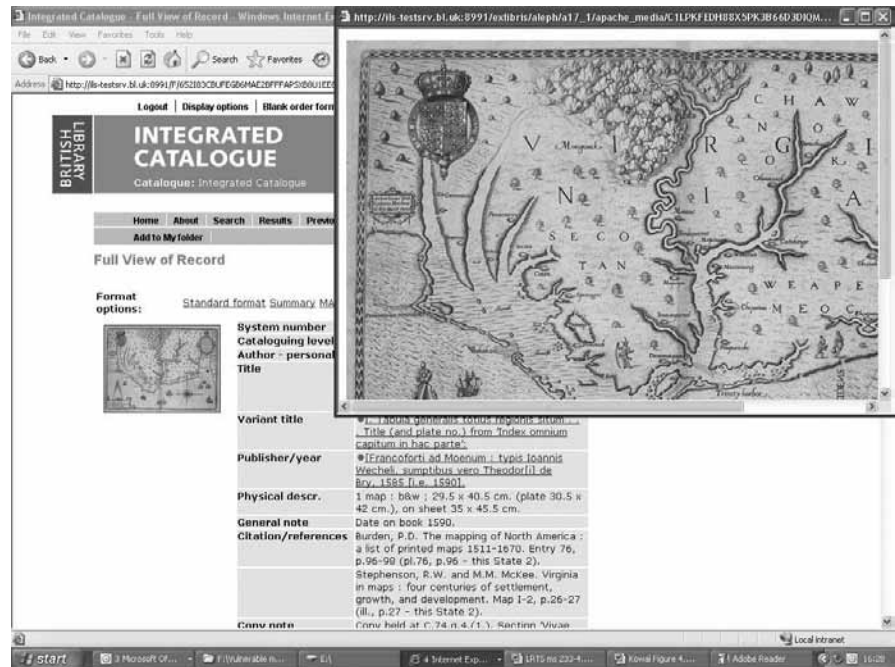
Figure 6. An Enlarged Access Image Generated by Clicking on the Thumbnail in the OPAC

from project to project, or if project tasks will be written into jobs.

Conclusion

This small digitization project at the BL was an opportunity to test how current cataloging codes and format standards can accommodate metadata and image capture within the ILS. The project successfully fulfilled the collection security needs of the organization while demonstrating that this approach can offer an improved product, thereby increasing the access and visibility of collection items to better meet the needs of researchers and providing the organization with data whose authenticity can be preserved and used in future systems. As opposed to purchasing a dedicated digital collection software suite or developing new websites that may or may not be found by library users, the collection items are integrated into the library catalog. The image, together with complete bibliographic information about both the map as an independent resource and its source book volume, may be retrieved with other library holdings in the OPAC.

The use of library system technology for creating and organizing metadata and making it searchable by users, and the MARC format with its flexibility and ability to handle differing levels of granularity and formats, is a powerful combination for handling digitized objects. The combination of established standards such as MARC21 and *AACR2* with the ILS, which operates in a similar way to many other ILSs, means that the approach described in this paper can be propagated to any library that uses these standards and has a comparable ILS to describe formerly "hidden" collection items.

References and Notes

1. Library of Congress, "On the Record: Report of the Library of Congress Working Group on the Future of Bibliographic Control" (Washington, D.C.: Library of Congress, 2008),

www.loc.gov/bibliographic-future/news/lcwg-ontherecord-jan08-final.pdf (accessed Aug. 18, 2008).

2. Kimberly C. Kowal and John Rhatigan, "The British Library's Vulnerable Collection Items Project," *LIBER Quarterly* 18, no. 2 (2008), http://liber.library.uu.nl (accessed Oct. 24, 2008).

3. Heidi G. Lerner and Seth Jerchower, "The Penn/Cambridge Genizah Fragment Project: Issues in Description, Access, and Reunification," *Cataloging & Classification Quarterly* 42, no. 1 (2006): 21–39; Ann Copeland et al., "Cataloging and Digitizing Ephemera: One Team's Experience with Pennsylvania German Broadsides and Fraktur," *Library Resources & Technical Services* 50, no. 3 (2006): 186–98; Jen Wolfe, "Digital Collection, the Next Generation: Transitioning to METS for a Science Fiction Digitization Project," *Against the Grain* 19, no. 1 (2007): 37–40; "NASA Internet Archive to Digitize Space Imagery," *Advanced Technology Libraries* 36, no. 10 (2007): 6; Susannah Benedetti, Annie Wu, and Sherman Hayes, "Art in a Medium-Sized University Library," *Library Resources & Technical Services* 48, no. 2 (2004): 144–54; James A. Bradley, "Zen and the Digital Collection Librarian," *Against the Grain* 19, no. 1 (2007): 32–36; Claire-Lise Bénaud, "Latin American and Iberian Posters: The Sam Slick Collection at the University of New Mexico," *Collection Management* 27, no. 2 (2002): 87–95; Walter E. Valero, Claudia A. Perry, and Tom Surpranant, "History on a Postcard," *netConnect* (Jan. 15, 2007), www.libraryjournal.com/article/CA6404150.html (accessed Nov. 15, 2008).

4. American Library Association, Map and Geography Round Table, "ALA MARGERT Map Scanning Registry," http://mapregistry.library.arizona.edu/cgi/index.pl (accessed Aug. 18, 2008).

5. Roy Tennant, "21st-Century Cataloging," *Library Journal* 123, no. 7 (1998): 30–31.

6. Jeffrey Beall, "Discrete Criteria for Selecting and Comparing Metadata Schemes," *Against the Grain* 19, no. 1 (Feb. 2007): 28–31.

7. Sarah Shatford Layne, "MARC Format for Medieval Manuscript Images," *Rare Books & Manuscripts Librarianship* 6, no. 1 (1991): 39–52.

8. Lerner and Jerchower, "The Penn/Cambridge Genizah Fragment Project"; Copeland et al., "Cataloging and Digitizing Ephemera."

9. *Anglo-American Cataloguing Rules*, 2nd ed., 2002 rev., 2005 update (Chicago: ALA; Ottawa: Canadian Library Association; London: Chartered Institute of Library and Information Professionals, 2005).

10. Christopher Cronin, "Metadata Provision and Standards Development at the Collaborative Digitization Program (CDP): A History," *First Monday* 13, no. 5 (May 2008), www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2085/1957 (accessed Nov. 15, 2008).

11. Nicholas Graham, "North Carolina Maps: Metadata Schema," University of North Carolina Libraries, unpublished document (May 9, 2008); Nicholas Graham, "MARC for Digital Project," e-mail to Kimberly C. Kowal, Aug. 1, 2008.

12. Christopher Cronin, "Cataloging Compound Digital Objects: Using METS for Digitized Sanborn Maps" (presentation at Introduction to Map Digitization, hosted by OCLC Preservation Centers, ALA Midwinter Meeting, Philadelphia, Jan. 12, 2008), http://libnet.colorado.edu/facultyprofiles/files/publications/libcroninc/OCLC%20ALA%20MW%20--%20Sanborn%20METS%20--%20No%20Notes.ppt (accessed Nov. 12, 2008).

13. See Terry Reese, "Low-Barrier MARC Record Generation from OAI-PMH repository Stores Using MarcEdit," on pages 57–70 of this issue.

14. Michaela Brenner and Peter Klein, "Discovering the Library with Google Earth," *Information Technology and Libraries* 27, no. 2 (2008): 32–36; Michaela Brenner, telephone conversation with Kimberly C. Kowal, Aug. 15, 2008.

15. T. Patrick McGlamery and F. Tyler Huffmann, "Building a Globally Distributed Historical Sheet Map Set," *University of Connecticut Center for Geographic Information and Analysis: Paper and Proceedings*, no. 2 (2007), http://digitalcommons.uconn.edu/cgi/viewcontent.cgi?article=1001&context=uccgia_papers (accessed Nov. 15, 2009).

16. Trudy Levy, "Digitizing in a Material world: A Digitization Symposium," *VRA Bulletin* 34, no. 3 (Fall 2007): 68–70.

17. British Library, Architecture and Development Department of the eIS Directorate, "E-strategy and Information Systems: Strategy for Architecture and Development," unpublished document (Apr. 2007).

18. Elizabeth U. Mangan, ed., *Cartographic Materials: A Manual of Interpretation for AACR2*, 2nd ed., 2002 Rev., 2005 Update (Chicago: ALA, 2006); Library of Congress, Network Development and Standards Office, "MARC 21 Format for Bibliographic Data, 1999 Edition Update No. 1 (October 2001) through Update No. 8," www.loc.gov/marc/bibliographic/ecbdhome.html (accessed Aug. 18, 2008).

19. *Anglo-American Cataloguing Rules*, 2nd ed., 2002 rev., 2005 update; Library of Congress, *Library of Congress Subject Headings*, 21st ed. (Washington D.C.: Library of Congress, 1998).

20. Library of Congress, Acquisitions and Bibliographic Access Directorate, "Core Data Set for 'Access Level' MARC/AACR Catalog Records," www.loc.gov/catdir/access/dataset_final.pdf (accessed Nov. 15, 2008).

21. Chris Martyn with Alan Danskin, "British Library Policy for Copy-Specific Information," unpublished document (British Library, Dec. 8, 2008).

22. Ibid.

23. The most up-to-date discussions take place on two maps electronic discussion lists: Maps and Air Photo Systems Forum (www.listserv.uga.edu/archives/maps-l.html) and MapHist, the Map History Discussion List (www.maphist.nl).

24. Karen Coyle, "The Library Catalog: Some Possible Futures," *The Journal of Academic Librarianship* 33, no. 3 (2007): 414–16.

Appendix. List of the Fields in the Template for the Analytic Bibliographic Record for the Map

FMT (The Map format was selected as the format in all cases, as it covers both maps and views.)

Leader (Type of record identified as Cartographic Material)

001—Record Control Number (The unique identifier for the record)

006—Fixed-Length Data Elements-Additional Material (A data element in this field indicating the materials' form denotes that the paper item cataloged has also been captured as a digital reproduction. This was used in preference to a second 007.)

007—Physical Description Fixed Field-General Information (A data element in this field, Category of material, indicates that the item is a map.)

008—Fixed-Length Data Elements (The following areas are used: date, place of publication, language, and type of cartographic material. For the latter, all are "g" to identify the item as "map bound within another work.")

034—Coded Cartographic Mathematical Data

040—Cataloging Source

100—Main Entry-Personal Name (These headings are subject to authority control.)

245—Title Statement (and 246)

255—Cartographic Mathematical Data (This cartographic materials-specific field, indicating scale, projection, and coordinates, was uniformly supplied as "scale not given." Scales were not generally supplied in a standard form on maps at the time, and most maps in the project that contained scale information expressed it in the form of a graphic scale or, in some cases, a verbal statement referring to scales no longer used, e.g., chains. Deciphering and translating either to a representative fraction in accordance with *AACR2* would have meant intensive labor producing only lukewarm results. Therefore the decision was made that the "scale not given" option for early cartographic materials would be applied.[*] This decision will be reviewed for phase two of the project, given the importance of scale. Geographic coordinates were not supplied for a similar reason. This too requires review, given the advantages of future potential display options to present materials in a geographical content. Both of these areas affect the contents of the correlated code field (034).)

260—Publication, Distribution, etc. (In most cases, this matched the date and place of publication for the book. In the case where there was a difference between the date printed on the map and that stated in the imprint of the book, the date on the map was recorded first, followed by the book imprint in brackets. A 500 note was created in explanation.)

300—Physical Description Area (The extent of the cartographic item was in all cases named as either "map" or "view." Most maps were printed in black and white, with less than 1 percent in color. Also in this area are listed the dimensions, i.e., height x width of the map, the plate, and the sheet. Measurements were rounded to the nearest half centimeter.)

510—Citation/reference note to published bibliographic descriptions, reviews, abstracts, or indexes (These citations provided additional information that could aid in deciphering the map, documenting the significance of the piece, and informing scholars that the map has been described extensively elsewhere. A maximum of three recent citations per record were referenced, with priority given to those works in English.)

583—Action Note: Preservation & Digitization Actions (This note records information about processing, reference, and preservation actions. The material specified was consistently "map.")

690—Local Subject (The records were united by the locally assigned "scanned maps and views.")

651—Subject Added Entry-Geographic Name (Although map records historically have used a dedicated subject system, the Map Library commenced using standard LCSH in 2004 with the move to Aleph. These headings are subject to authority control.)

---

[*] Elizabeth Mangan, ed., *Cartographic Materials: A Manual of Interpretation for AACR2*, 2002 Rev., 2005 Update, 2nd ed. (Chicago: ALA, 2006), Appendix G.2.