Beyond the Back Room: The Role of Metadata and Catalog Librarians in Digital Humanities

Lisa M. McFall
Hamilton College, lmcfall@hamilton.edu

Follow this and additional works at: https://digitalcommons.hamilton.edu/chapters
Part of the Cataloging and Metadata Commons, and the Digital Humanities Commons

This chapter appeared on pages 21-43 of:
Find this book in a library.

Citation Information
McFall, Lisa M., "Beyond the Back Room: The Role of Metadata and Catalog Librarians in Digital Humanities" (2015). Hamilton Digital Commons.
https://digitalcommons.hamilton.edu/chapters/1

This work is made available by Hamilton College for educational and research purposes under a Creative Commons BY-NC-ND 4.0 license. For more information, visit http://digitalcommons.hamilton.edu/about.html or contact digitalcommons@hamilton.edu.
2015

Supporting Digital Humanities for Knowledge Acquisition in Modern Libraries

Lisa M. McFall
lmcfall@hamilton.edu

Follow this and additional works at: http://digitalcommons.hamilton.edu/chapters

Recommended Citation

This Book Chapter is brought to you for free and open access by Hamilton Digital Commons. It has been accepted for inclusion in Chapters by an authorized administrator of Hamilton Digital Commons.
Supporting Digital Humanities for Knowledge Acquisition in Modern Libraries

Kathleen L. Sacco
State University of New York at Fredonia, USA

Scott S. Richmond
State University of New York at Fredonia, USA

Sara Parme
State University of New York at Fredonia, USA

Kerrie Fergen Wilkes
State University of New York at Fredonia, USA

A volume in the Advances in Library and Information Science (ALIS) Book Series
Chapter 2
Beyond the Back Room: The Role of Metadata and Catalog Librarians in Digital Humanities

Lisa M. McFall
Hamilton College, USA

ABSTRACT
Many professors and students are surprised to learn about the variety of skills metadata and catalog librarians possess that can benefit their digital humanities projects. Metadata and catalog librarians already have experience in areas such as developing project metadata schemas and controlled vocabularies, as well as providing suggestions for project organization, and have the basic abilities needed to support learning new skills such as XML and TEI. This chapter will offer perspectives on how a metadata and catalog librarian can contribute to digital humanities projects. A case study focusing on the involvement of the Metadata and Catalog Librarian with the Digital Humanities Initiative (DHi) at Hamilton College will be used to illustrate one example of how to become successfully involved in digital humanities research on campus.

INTRODUCTION
Public services librarians have been considered the face of the library for decades (Rothstein, 1953), and the interaction of faculty members and these librarians has been the topic of many books and articles (Dilmore, 1996; Raspa & Ward, 2000; Kraat, 2005). Faculty members frequently interact with public services librarians, whether it is for assistance in locating items for their own research, arranging a course lecture on subject-specific resources, or pointing students towards the services of reference librarians when beginning a research paper. Ducas and Michaud-Oystryk (2003) divided the interactions of faculty members and primarily public services librarians into the categories of teaching/instruction, information services, information technology, research, and collections. Through these communications, faculty members become accustomed to the services that reference and research librarians offer, but many are surprised to learn about the skill sets that technical services librarians possess that can assist them in...
Beyond the Back Room

their work. While reference, instruction, and research librarians are very necessary partners in digital humanities projects, it is becoming apparent that the unique expertise of metadata and catalog librarians has a new and prominent place in supporting digital humanities projects on campus.

Much has been written on the topic of digital humanities, with articles appearing in discipline-specific journals, as well as in journals exclusively dedicated to digital humanities research. Digital Humanities Quarterly (2007-present) and the Journal of Digital Humanities (2011-present) offer an outlet for articles and reviews on a variety of aspects of digital humanities projects. Publications such as Gibb’s 2011 article about critical discourse in digital humanities provide an entry point for those who are new to this field.

The roles that librarians can (and do) play in digital humanities is a recent field of interest, as exemplified through an informal survey conducted by the Association of College & Research Libraries (ACRL) Digital Humanities Discussion Group, and as addressed in an article by Coble (2012) that offers suggestions on how to evaluate the contributions of librarians who work in digital humanities. A few short publications exist that focus on how to get involved in and learn about digital humanities, notably Spiro’s 2011 article intended for a general audience and Adams and Gunn’s 2012 article geared towards librarians; however, a practical guide for metadata and catalog librarians who are “doing” digital humanities is notably absent from the literature.

Metadata and catalog librarians can be involved in digital humanities projects in a variety of ways. Their experience in evaluating, implementing and customizing metadata schemas for library collections can be reapplied to working one-on-one with faculty members on this aspect of building their digital humanities project. Additionally, metadata and catalog librarians can work directly with faculty and students to help them learn and develop a variety of technical skills needed to work in the digital humanities. To accomplish this, a deeper understanding of the various parts of digital humanities research and new technical abilities, such as in working with XML and TEI, may need to be developed.

The purpose of this chapter is to offer perspectives on how a metadata and catalog librarian can contribute to digital humanities projects, particularly on campuses that do not have the means to have a metadata position dedicated exclusively to supporting digital humanities. Topics addressed will include the importance of collaboration on digital humanities projects, useful skills that transfer from traditional metadata and catalog librarian positions to the work that is done in supporting digital humanities, recommendations for additional skill areas to investigate and where to acquire these skills, and suggestions for how best to work with faculty and students on digital humanities projects. A case study found in this chapter focuses on how the Metadata and Catalog Librarian at Hamilton College developed the additional role as the metadata consultant for the Digital Humanities Initiative and will provide one example of how “traditional” metadata and catalog librarians can successfully expand their role into supporting digital humanities projects, while also demonstrating to collaborators the important skills that technical services librarians can bring to supporting this relatively young field.

COLLABORATION IS KEY: DEVELOPING PARTNERSHIPS ON CAMPUS

Digital humanities projects truly encompass a team effort and rely on the involvement of many people with a wide variety of skills. Faculty members are often the generator of the initial project idea and data, but to build their project in the digital realm, they require support from various partners. The information technology department frequently helps to support these projects in the form of server support, database management, and website design and structure. Public services librarians and staff in the library may
help through providing assistance in locating materials to be added to the archive and providing course support for classes that incorporate these digital humanities projects into course design. Through the skills developed while working with the library’s digital collections, metadata and catalog librarians fit well into the support structure for digital humanities research and provide a unique set of skills to assist in digital humanities projects.

**Contributions of Metadata and Catalog Librarians**

As libraries and the collections that they make available have evolved, new roles beyond the traditional cataloging of printed materials have developed for metadata and catalog librarians. Boydston and Leysen explored the current roles and responsibilities of catalog (and metadata) librarians in their 2014 article about the evolving work of ARL cataloger librarian responsibilities. These librarians are now frequently involved in providing access to institutional repositories and the library’s digital collections, which often include collections of items in a wide variety of formats such as video and audio recordings, visual materials such as postcards and photographs, and printed materials such as books, pamphlets and other ephemera. It is through this work that metadata and catalog librarians have gained a broader sense of the application of various metadata schema and database systems that can be valuable to those working on digital humanities projects.

**Metadata and System Knowledge**

As described later in this chapter, a wide variety of generic and specific metadata schemas exist. Many metadata and catalog librarians are familiar with Dublin Core, which is commonly used in digital collection management software such as OCLC’s CONTENTdm. Some technical services librarians may also have had the opportunity to explore other systems, such as the open source Fedora/Islandora, Digital Commons by bepress, and customized, homegrown repositories, which frequently provide some additional flexibility in the use of other metadata schema. Each system may, however, have certain limitations or requirements of the metadata schema that can be used in it. Despite these differences, the principles of working with a database system that is created to house digital images and both born-digital objects and digital surrogates with their associated metadata remains the same and the basic knowledge gained from working with these metadata schema in various systems is transferrable to working with metadata for digital humanities projects. Few other people working in the field of digital humanities are likely to possess the same in-depth, first-hand experience of the integration of metadata schemas into a database structure as metadata and catalog librarians.

Metadata and catalog librarians who work with digital collections are generally open to innovations and creative problem solving to accommodate the specialized needs of digital objects and their metadata, including adapting familiar metadata schemas and controlled vocabularies. Unlike traditional cataloging in MARC21 where strict rules are employed on how metadata can be entered into the schema, there is an additional level of flexibility inherent in non-MARC metadata schemas, mostly due to the wide variety of items that are digitized and have a need for specialized descriptive metadata. Faculty projects in digital humanities often have unique needs for their metadata that do not fit neatly into the predefined elements of established metadata schemas. With their knowledge of how metadata schemas are used by systems for searching and their previous work in using non-MARC metadata schemas, metadata and catalog librarians can find ways to make the project-specific metadata work in the underlying database system used as the framework for digital humanities projects.
Beyond the Metadata: Other Contributions and Considerations for Metadata and Catalog Librarians

Along with working with a variety of metadata schemas and knowledge of system use of metadata, metadata and catalog librarians may also possess familiarity with the process of and specifications for scanning images, knowledge of best practices for digital object preservation, and how to create project documentation, such as a data dictionary to record local metadata practices. All of these skills are also applicable to the field of digital humanities. This combination of wisdom and experience from working with the digital collections of the library makes these librarians ideal collaborators on digital humanities projects.

While metadata and catalog librarians may be accustomed to creating all of the metadata for the library’s digital collections on their own, it is unlikely that they will be able to do the same for faculty digital humanities projects, mainly due to time constraints but also due to the unique subject expertise faculty bring to their digital projects. Once the amount of time a metadata and catalog librarian can devote to digital humanities projects is ascertained, a specific plan can be designed for how the metadata needs of digital humanities projects can be met. This provides an opportunity for technical services librarians to take a leadership role in training faculty and students to be the primary creators of project metadata. Successful partnerships with faculty and students on campus allow metadata and catalog librarians to play an active role in digital humanities projects while balancing this work with the rest of their responsibilities.

Working with Faculty

Faculty members use a variety of resources and services in college libraries, and often think that they are familiar with the people who work in the library. Frequently, however, they are not as aware of the tasks that happen in the “back room” of the library, and are therefore unfamiliar with the skills that technical services librarians have to assist them in projects. Metadata and catalog librarians can use their specific skills to support and contribute to faculty projects in a variety of ways.

While faculty members are usually very familiar with databases, electronic journals, and websites that are used in their particular discipline, they may not understand the inner workings of how these resources utilize metadata to create the functionality upon which they rely. Faculty members are quite familiar with their own research project and the goals they see for the project -- in the traditional sense of publishing -- but they may not have envisioned exactly how users will access and navigate through the objects in their collection. This can make it challenging for them to think about how to structure their metadata and what information should be recorded. Metadata and catalog librarians, on the other hand, regularly think about how users will access the cataloged materials; this knowledge should be shared with faculty developing projects in the digital humanities. Once it has been demonstrated how this impacts the project, it may be possible for the faculty member to think about the site structure in a new way.

It may also be helpful to ask questions of faculty to get them to think about the user scenarios on their own. These questions may include:

- Who is the primary target audience for your collection?
- How do you envision experts in this subject interacting with the collection?
- How will students and those who are non-experts use your site?
Beyond the Back Room

• If you were to stumble upon this site, how would you expect to find resources and what are some of the key categories and characteristics by which you would expect to be able to search?

Questions such as these allow faculty members to begin thinking about their collections in a user-focused manner. It may also be beneficial to talk to the faculty member about other sites and projects that are seen as inspiration for their own research and to go through the sites together, talking about what the faculty member likes and dislikes about each of them. As part of this exploration, it may be helpful for the metadata and catalog librarian to point out the role that metadata has in the navigation of objects found in these projects. Once this process is underway, the metadata and catalog librarian can offer suggestions for how to tailor a custom metadata schema to the digital collection that the faculty member is seeking to build.

After the metadata schema has been developed with input from the faculty member, it is also advisable for the librarian to sit down with the faculty member to work together in creating the metadata for a sample object or two. Demonstrating this process to faculty members will allow them to develop a clearer understanding of what is expected when completing metadata on their own for other objects in the collection. Completed sample metadata records can be used as a point of reference for faculty and students to follow as examples of how to properly complete metadata when the metadata librarian is not available to work one-on-one during the metadata creation.

Working with Students

Metadata and catalog librarians can work with undergraduate students to support the growth of digital humanities scholarship on campus and foster the students’ own technical skills in a variety of ways. Trained students are excellent collaborators for faculty members, and can assist with a variety of tasks to support their projects, such as helping to digitize objects and creating metadata for items in the collection.

Students may be able to assist with scanning documents and photographs that are to be added to the faculty member’s project. The metadata and catalog librarian can teach students about scanning specifications for the project, setting guidelines for scanning resolution and cropping. They can also develop a file naming schema to be used by the students and faculty to organize the project files. Once an object has been scanned, students can then record the file name in the metadata record, along with other technical metadata such as file size and type (choosing from a selected list of MIME Types). If a document undergoes OCR (Optical Character Recognition), a student can proofread the document to ensure that it is correct. Depending on project goals, meticulous students may also be entrusted with encoding some basic TEI elements and formatting.

In addition to processing the physical items and files, students are also capable of entering descriptive metadata. Most projects can easily accommodate student involvement in the creation of metadata based on what types of objects the project contains and the level of confidence the faculty member has in their student’s abilities. A faculty member and the metadata and catalog librarian can work together to determine which specific metadata is appropriate for students to complete. For example, if a subset of keywords is created, students may be able to apply those keywords to objects or enter information such as the object’s title, creator/contributor, and relevant dates. Training for the student, as well as detailed documentation for how to correctly enter the metadata, can be provided by the metadata and catalog librarian.
Beyond the Back Room

Tools for Supporting Collaborative Metadata Creation

Faculty and their student collaborators can both benefit from documentation related to inputting metadata. Documentation should include information about selected lists of controlled vocabularies (both locally-defined and nationally recognized) for specific fields, how to input information such as the proper standardized format of entering dates and names, and an explanation of the type of data that is entered into each field. While it is not uncommon in libraries for only one person to be the creator of metadata for any given object or collection, there are multiple people involved in each digital humanities project who are likely to create metadata. Clear guidelines can ensure that the metadata that is entered for each project (and across projects in the same database) is consistent.

This documentation can be presented in a variety of ways. To save time on future projects, information that is non-project specific and expected of items in all faculty digital humanities projects can be generalized and put on a wiki devoted to metadata for all digital humanities projects. Included in this can be a detailed data dictionary that provides key information such as specific instructions on how to record metadata (e.g., dates should be recorded YYYY-MM-DD, names should be entered LastName, FirstName, etc.) and examples of the correctly recorded metadata (e.g., 2014-01-01, Smith, Jane R., etc.). Links to external controlled vocabularies, as well as lists for valid internally customized controlled vocabularies can also be maintained on this wiki. The wiki can also be an ideal place to record project-related decisions related to the metadata and other project aspects. Having one key place to track this information can make it much easier to support projects that may take several years – and potentially several personnel changes – to complete.

In addition to a data dictionary defining the input conventions for fields that are used across multiple projects, it is also important to provide guidance for project-specific fields. If spreadsheets are used for inputting project metadata, comments can be added to each column header with a brief description of how the metadata should be added, along with an example of the metadata in the properly formatted structure. This method allows key information to be embedded within the document and can serve as a quick guide of sorts for those working on the spreadsheet. Implementing such detailed and robust support will allow faculty members and students to have a degree of autonomy as they create the metadata, while providing them with the tools to make the metadata as standardized as possible.

CASE STUDY: THE DIGITAL HUMANITIES INITIATIVE

Hamilton College formed the Digital Humanities Initiative (DHi) in 2009. The purpose of DHi was to create “an incubator hub where new media and computing technologies promote inter- and multi-disciplinary humanities-based scholarship, ultimately leading to the creation of knowledge and curricular innovation” (The Digital Humanities initiative (DHi) at Hamilton College, 2010, p. 1). DHi pioneered the effort to develop a method for supporting digital humanities on a small, liberal arts college campus where there was not a dedicated center for digital humanities with full-time positions available for supporting such work. In 2010 and 2013, DHi received Andrew W. Mellon Foundation grants to further develop and support its efforts.

DHi formed the Collection Development Team in 2011 and included members of the library, information technology department, and two unique positions funded through the Dean of Faculty. The team was conceived to build collaborations in support of faculty research, while making the resulting collections
Beyond the Back Room

of objects, metadata, and interpretation available through a web-based platform. The diverse membership and skills of the Collection Development Team allows it to provide a broad variety of assistance to faculty project directors, including developing custom metadata schemas, ingesting collections into a project database, and creating customized website designs and user interfaces. Members of the team include the Co-Directors of DHi (one an educational technologist on campus and the other a tenure faculty member), the Lead Designer and Software Engineer, the Unix/HPC Systems Administrator, the library’s Information Systems Specialist, and the Metadata and Catalog Librarian from the technical services department of the library. In addition to individual responsibilities related to project development, members of the Collection Development Team work collaboratively to brainstorm creative solutions that support innovation in a faculty member’s research and analysis. Over the course of nearly four years, the team has worked together on a number of tasks, including deciding upon a technological infrastructure, developing a project workflow and best practices, as well as developing a customizable metadata schema and the appropriate ingest scripts.

Building a Digital Humanities Project

In determining the best approaches to software and design, members of DHi and the library reviewed a number of different technological solutions. The decision was made to use a Fedora Commons Repository hosted at Hamilton College, with Islandora as a “middleware” solution between Fedora and the Drupal-based front end. This configuration allowed for the use of MODS in project metadata (in addition to Fedora’s automatically-generated Dublin Core records), providing the ability to adapt the MODS <extension> element for project-specific fields.

Over the course of the first year of working with several faculty members on their projects, DHi developed a standardized, multi-step approach to supporting new faculty digital humanities projects. Faculty members are guided through several steps to ensure that the final project exhibits their research as envisioned and that there are connections to a course or courses taught at Hamilton College. After the project has first been vetted by the DHi co-directors, the process continues with a meeting involving the Collection Development Team and the faculty member, with the goal of developing a better understanding of the faculty member’s vision and the technological requirements of the project. Once the team has decided that the necessary technology to support the project either already exists or can be developed and the project is approved, specific meetings with various subgroups are planned to further develop a comprehensive understanding of the specific goals and needs of the project. Subgroup meeting topics include project site design and the development of the metadata schema. It is in the latter of these that the metadata and catalog librarian plays an integral part.

The Role of the Metadata and Catalog Librarian

The metadata schema meetings involve the faculty project scholar, the metadata and catalog librarian, the DHi co-director and sometimes the faculty member’s student assistant(s). The purpose of these meetings is to allow the faculty member to coordinate with the metadata and catalog librarian about the development of a metadata schema specific to the project. As part of this initial metadata meeting, the faculty member explains how they envision visitors to the digital archive interacting with the site. As described above in the section titled “Working with Faculty,” the metadata and catalog librarian talks with the faculty member about digital sites that they use and how the metadata facilitates certain layers
of that interaction. This then evolves into discussing the role that metadata will play in the way people interact with the faculty member’s site, along with the manner in which the website and repository will use it. Many faculty members are amazed to see how metadata functions in an archive, from the way it is used to implement advanced searches to its use in faceted searches and possible implications for integrating a map of resources into the site. Once these connections are explained, faculty members begin to understand why the metadata for the project is important, and discussion moves to the goals for the faculty member’s project.

Project-specific discussion revolves around topics such as the target audience(s) of the archive, the intended use of the archive (research, classroom setting, etc.), and how the development of the metadata can support these tasks. The faculty member is asked about how they imagine interacting with their own archive, what they hope a visitor to the site will discover, and how the visitor will navigate through the collection. By developing use cases for the archive, it is possible for both the faculty member and the metadata and catalog librarian to gain a better understanding of the steps necessary to complete these goals and how a robust, customized metadata schema can help support the desired outcome.

The Development of the Project’s Metadata Design

The template spreadsheet, which illustrates the types of metadata that are required for all DHi projects, is used as the basis for the design of each project’s customized metadata spreadsheet. Required metadata that is standardized across all DHi projects includes:

- Title,
- Creator/contributor (if applicable),
- Language used in item,
- Tags/keywords,
- Date of digital object creation and any other relevant dates (such as original date of creation of the counterpart analog item),
- Publisher (if applicable),
- File size,
- Standardized file name,
- Submitter of object,
- Local genre (using a customized controlled vocabulary),
- Copyright/rights information,
- Type of resource (using the Art and Architecture Thesaurus),
- MIME Type, and
- Library of Congress Subject Headings.

The template spreadsheet is divided into three categories: (1) fields that are completed by faculty/students that are required for all DHi projects, (2) administrative and descriptive metadata fields that are completed by the metadata and catalog librarian that require specialized controlled vocabularies (such as AAT and LCSH), and (3) custom fields for the individual faculty member’s project.

Relationships between objects in the project are also documented. The columns of the spreadsheet are sorted and color-coded according to the three categories above, which makes it easier for faculty, students, and the metadata and catalog librarian to determine at a glance who is responsible for complet-
Beyond the Back Room

ing each column of metadata. A separate worksheet is then set up for each type of file represented in the project (e.g., photographs, newspaper articles, letters, etc.) due to the need to record different types of metadata for different types of objects. Providing a separate worksheet for each type of object allows the user to be very specific in the metadata they record, permitting them to record a letter’s sender and recipient with their particular assigned roles of sender and recipient, for example.

Additionally, a wiki with information about how to complete the metadata for each required field was created as a resource for faculty project scholars and student fellows. This wiki includes instructions on how to enter metadata into the various fields (e.g., using sentence case for titles, entering the date YYYY-MM-DD, etc.), links to relevant controlled vocabularies, and examples of how to complete the metadata entry.1

Working with Faculty and Students

Once faculty members have received an overview of required metadata, a discussion begins about project-specific fields the scholar would like to incorporate. Suggestions are offered by the metadata and catalog librarian for metadata that the faculty member might want to use in the collection, including some categories that the faculty member may not have considered previously. As the faculty member begins to see how metadata can serve the purpose of developing layers of navigation and interpretation on the site, their enthusiasm for working on creating metadata may increase substantially. Through these discussions, the metadata and catalog librarian becomes an integral part of the faculty member’s project and often the faculty member is able to understand not only the value of metadata, but the skills and insights that the librarian can offer to the project as well.

The metadata and catalog librarian works to customize the general project spreadsheet once the initial meeting and metadata brainstorming session is completed. Columns specific to the faculty member’s project are added to the appropriate color-coded section of the spreadsheet. These customized fields can vary widely from project to project, where past examples from DHi projects include fields such as race, sex, prison security level, and protection status of historical sites. Key guidelines and a brief example of how to enter the metadata are included as a comment in all column headers for the project. Although faculty and students are able to reference the wiki for information about the columns required for all DHi projects, the comments in the spreadsheet are the primary resource for how to complete the project-specific metadata fields. Links are included in the comments for any necessary external controlled vocabularies.

Once the metadata spreadsheet has been updated for the faculty member’s needs, a second meeting is held to review the revised spreadsheet and collaboratively create the metadata for a sample object or two. This metadata serves as an example for faculty members as they complete the metadata for the rest of the project. The faculty project scholar is then asked to complete the metadata for 5-10 items on their own and after these are completed, the metadata and catalog librarian reviews the records and offers suggestions for improvement. Once the librarian is confident in the faculty member’s ability to complete the metadata on their own, the faculty member is released to complete the rest of the project metadata. The metadata and catalog librarian maintains contact with the faculty member, checking on their progress and answering any questions as they arise.

Student participation is crucial to the development of the project. The DHi co-director trains students in scanning techniques and specifications developed by the library’s information systems specialist. While scanning, the students apply a file naming system designed by the metadata and catalog librarian to organize and identify project files. Students also complete basic project metadata including associ-
ated dates (following proper formatting and input standards), file size, language, and submitter. Other descriptive metadata, such as keywords/tags, title, and creator/contributor, may also be completed by the student at the discretion of the faculty project scholar.

Depending on the number of items in the collection, the number of customized fields, and the amount of time the faculty member and student have to dedicate to the project’s completion, the faculty member’s metadata may be finished in a matter of weeks, or over the course of several months. The metadata and catalog librarian is available to provide assistance as necessary during this process.

**Final Stages of the Metadata Processing**

After the scholar and student have completed their share of the metadata, the librarian reviews it -- checking for the correct use of controlled vocabularies and verifying that the metadata has been entered as required in the project guidelines -- and then completes the fields that remain. The library systems specialist on the Collection Development Team, with input from the metadata and catalog librarian, customizes and runs scripts that convert the final metadata spreadsheet into MODS XML files. Because DHi projects
utilize MODS, most of the faculty member’s customized metadata fields in the spreadsheet are mapped
to the MODS <extension> element. This allows the metadata to be robust and flexible while remaining
within the specifications for the other MODS elements.

One MODS XML file is created for each object that will be added to the repository. The metadata
and catalog librarian performs a quality control check, at which point the objects and their associ-
ated metadata are ready to be ingested into Fedora by the library information systems specialist. The
site is then ready for design and theming efforts, which incorporate the metadata into each object’s
user interfaces. The librarian continues to contribute to the project by reviewing the integration of
the metadata into the website and running tests to be sure that the metadata is functioning correctly
in the database.

DHi’s project process is designed to provide faculty project scholars with the support they need
to bring their projects to fulfillment. The involvement of the metadata and catalog librarian in this
process ensures that the design and quality of the metadata supports the robust needs of a digital
humanities project. By working directly with them to develop a metadata schema, faculty see first-
hand the valuable contributions metadata and catalog librarians can make to support their project
and research.

RECYCLED SKILLS IN THE DIGITAL HUMANITIES

Beginning to work on projects in the digital humanities can sometimes feel daunting due to the amount
of new skills that seem necessary. It is important to remember that metadata and catalog librarians al-
ready possess some valuable skills from cataloging that transfer well to working on digital humanities
projects. Some of the most widely-used metadata schemas include Dublin Core (DC), the Metadata
Object Description Schema (MODS), Cataloging Cultural Objects (CCO), and the Visual Resources
Association Core (VRA Core). These schemas have been developed to accommodate digital objects that
are now preserved by institutions. Depending on the needs of the project, any of these schemas can be
adopted for use in a digital humanities archive.

Cataloging Cultural Objects

Cataloging Cultural Objects, or CCO, follows the principle of organizing objects around the concept
of a Work Record (Visual Resources Association, “10 Key Concepts,” n.d.) and is intended to serve
as a metadata schema for institutions, including museums, who catalog cultural objects and their
digital surrogates. CCO has several elements that are required for all objects, including work type,
title, creator, measurements, materials and techniques, and display date (Baca, M., Harpring, P.,
Lanzi, E., McRae, L., and Whiteside, A., 2006, p. 44-45). In addition to being available to purchase
in a print edition, the book Cataloging Cultural Objects: A guide to describing cultural works and
their images is available as a free PDF download and serves as a useful resource that includes both
the guidelines and relevant examples (Visual Resources Association, “CCO news and events,” n.d.).
Digital humanities collections working with different versions of the same item may find inspiration
in the organizing principles of CCO.
Dublin Core

Dublin Core originated in 1995 from a joint workshop of librarians and others looking to design a simplified description schema for electronic resources (Weibel & Lagoze, 1997, p. 176) and was originally designed with fifteen elements for use, which included the following: contributor, creator, date, description, identifier, subject, and title (DCMI, “Dublin Core Metadata Element Set, Version 1.1, 2012). The development of Qualified Dublin Core shortly after the release of the original Dublin Core was aimed at providing a more robust metadata schema for those implementers who wanted the ability to further qualify their metadata. Refinements included adding the ability to specifically denote an alternative title, a table of contents, and specific types of dates such as date created, date issued and date modified. Dublin Core is one of the most ubiquitous metadata standards and most systems can accommodate metadata in Dublin Core, making this an appealing schema for those interested in metadata migration.

Metadata Object Description Schema

The Metadata Object Description Schema, or MODS, is a descriptive metadata schema that was developed by the Library of Congress’s Network Development and MARC Standards Office in the early 2000s (Guenther, 2003, p. 138). It is used within an XML structure (see the “Useful Skills in Digital Humanities” section of this chapter for a description of XML) and serves as an alternative to metadata schemas with a very limited number of elements, such as Dublin Core, and very complicated metadata schemas such as MARC21 (Guenther, p. 139). The flexibility of MODS is also furthered by the availability of the <extension> element, which allows for those developing local uses of the MODS schema to expand their element set outside of that built into MODS. This can be applied by using the <extension> element to record metadata in other established metadata schemas (such as CCO or other subject-specific schemas) or to connect with a locally developed and maintained namespace of metadata elements that are institution-based or project-based. One example of the use of this <extension> element in a digital humanities project is demonstrated in the DHi case study discussed earlier in this chapter.

Visual Resources Association Core

Currently in its fourth iteration, the Visual Resources Association Core, better known as VRA Core, is a metadata schema designed to take into account the specific needs of describing a collection of visual materials and their digital surrogates. Designed to be used with CCO, VRA Core 4 indicates the relationships between works and the corresponding images that document them (Visual Resources Association, “Frequently asked questions,” n.d.). VRA Core 4 is hosted by the Library of Congress and has been translated from English into Italian and Greek (Library of Congress, 2014). The main elements for VRA Core include identification of the work, collection or image, the agent, date, location, material, stylePeriod, and technique (Library of Congress, 2007). Like CCO, this schema may be of interest to those working on digital humanities projects examining different iterations of visual objects.
Some new skills may need to be acquired to adequately support digital humanities projects. Many metadata and catalog librarians are accustomed to working in a database structure that hides some of the encoding needs of metadata and may need to expand upon their technical skills that allow them to work directly with the metadata; knowledge of XML makes this possible. TEI and an understanding of geospatial metadata are other useful skills that metadata and catalog librarians who support faculty work in digital humanities may need to develop.

**XML**

In a library setting, metadata and catalog librarians may frequently work with metadata in integrated library systems or digital collection software such as CONTENTdm. These programs often have their own method of inputting metadata, ranging from a form with labeled fields to an Excel spreadsheet for uploading metadata. While some systems that are used to house digital humanities projects may have similar options, metadata and catalog librarians may find the need to use XML to format metadata directly.

EXtensible Markup Language, or XML, is a flexible method of encoding data using customizable tags. It can be used in conjunction with controlled vocabularies and encoding standards such as MODS. XML records declare a namespace at the start and feature tags that serve as field names. Each element has an opening and closing tag.

XML can be edited using a variety of free or commercial programs. Although a simple text editor allows a cataloger to work with the tags and encoded data, programs such as the popular oXygen that are specifically designed to edit XML documents may offer more advanced functionality, proving easier for regular use. oXygen allows users to work with both XML and Extensible Stylesheet Language Transformations (XSLT) files, and provides a color-coded view of fields and attributes to make navigating the XML files easier.

A variety of tools exist for learning more about XML. The popular subscription-based training site Lynda.com offers a class titled “XML Essential Training” that provides a technical overview of what XML is and how to use it. Many other websites, such as W3schools.com, offer free tutorials and exercises for learning XML.
TEI

The Text Encoding Initiative, or TEI, is a method of encoding text that provides layers of searching and interpretation by using XML for markup. A wide variety of TEI tags exist, but a subset customized at the project level may be best to provide a tighter control of the elements that are used in a given project. Sample tags include <abstract>, <author>, <edition>, <keywords>, <persNam> (i.e., personal name), and <title>. A full set of acceptable tags can be found on the TEI Consortium (TEI-C) site (Text Encoding Initiative, “Appendix C Elements,” 2014). Those starting a TEI project for the first time may be interested in exploring TEI Lite, a “specific customization of the TEI tagset designed to meet ‘90% of the needs of 90% of the TEI user community’” (TEI, “TEI Lite,” n.d.).

Many resources exist to assist in learning TEI. The website maintained by TEI-C provides extensive resources to support learning about and implementing TEI. A page called “Learn the TEI” on the TEI-C site provides a straightforward overview including an introduction to TEI, a calendar of TEI training events, self-instruction tutorials, an extensive bibliography of resources on TEI, and an introduction to the essentials of XML, including information about its structure and attributes (TEI, “Learn the TEI,” n.d.). Many institutions also offer their own list of suggested guidelines and information about their own implementations of TEI that can be used as examples for establishing TEI practices. Additionally, a site called TEI By Example offers a free validation service to check TEI code and return a list of errors (TEI By Example, n.d.).

Geospatial Metadata

While many metadata and catalog librarians may have worked with geospatial metadata in the context of the library’s digital collections, geospatial metadata is being used in new and exciting ways in digital humanities projects. It is important that metadata and catalog librarians develop an understanding of how metadata is used in the project and tools that exist for working with geospatial metadata. Metadata schemas such as MODS offer elements and attributes for including coordinate data (<subject> <cartographics> <coordinates>), as well as geographical hierarchies (<subject> <hierarchicalGeographic>). Information about geospatial metadata and incorporating it into projects can be found through organizations including the Federal Geographic Data Committee (FGDC). The FGDC website offers a resources page providing suggestions for those looking at creating geospatial metadata and tools to work with geospatial metadata (Federal Geographic Data Committee, 2014).

LEARNING NEW SKILLS

It may seem daunting to become an “expert” in the new skills that are often required of metadata and catalog librarians who work with digital humanities projects. In addition to the subject-specific resources mentioned above, there are other resources with which to learn about the digital humanities more broadly, to see examples of digital humanities research in action, and to learn about new tools and cutting-edge developments that may benefit digital humanities projects on campus.
THATcamps

The Humanities and Technology Camp (THATcamp) is an “unconference” that serves as an informal meeting of humanists and technologists at a wide range of skill levels, convening to share information related to the intersection of technology and humanities (THATcamp, n.d.). THATcamps exist all around the world and last from one to three days, covering a variety of topics including crowdsourcing, text mining, teaching digital humanities, and Geographic Information Systems (GIS). A list of registered THATcamps can be found on the THATcamp website at: http://thatcamp.org/camps/.

Conferences and Training Sessions

As the field of digital humanities has grown, conferences focused entirely on digital humanities have multiplied. The Alliance of Digital Humanities Organizations (ADHO) sponsors an annual, multi-day international conference focused on digital humanities. Although this is not specific to metadata librarians, previous conferences have included sessions such as “MESA and ARC, developing disciplinary metadata requirements in a multidisciplinary context,” “Optimized platform for capturing metadata of historical correspondences,” and “Using computer vision to improve image metadata” (University of Nebraska, Lincoln, 2013; “Digital humanities, Lausanne – Switzerland, ’14”, n.d.).


The Digital Humanities Summer Institute (DHSI) offers intensive, one-week classes on specific topics related to digital humanities. Classes have been offered on TEI, digitization, XSLT and other potential topics of interest to metadata and catalog librarians involved in the digital humanities (Digital Humanities Summer Institute, n.d.).

code4lib is a community that includes members from a variety of professional areas including “hackers, designers, architects, curators, catalogers, artists and instigators from around the world” (code4lib, “About,” n.d.) that hosts an annual conference. Sample presentations that metadata and catalog librarians may have interest in include papers titled “Helping Google (and scholars, researchers, educators, & the public) find archival audio,” Making your digital objects embeddable around the web,” “EAD without XSLT: A practical new approach to web-based finding aids,” “ALL TEH [sic] METADATAS! or how we use RDF to keep all of the digital object metadata formats thrown at us; “All teh [sic] Metadatas Re-revisited,” and “Opinionated metadata (OM): Bringing a bit of sanity to the world of XML metadata” (code4lib, “2015 conference schedule”; code4lib, “Code4Lib 2013 schedule”; code4lib, “Code4Lib 2012 schedule”; code4lib, “Code4Lib 2011 schedule”).
Webinars and Online Classes

For those who may have trouble funding traveling to attend conferences and THATcamps, the online environment offers some possible alternatives. Professional organizations, often under the umbrella of ALA, frequently offer online classes and webinars geared towards helping metadata and catalog librarians expand their skills. LITA has offered a web course titled “Getting started with GIS,” taught by Eva Dodsworth, author of the LITA publication *Getting started with GIS: A LITA guide* which provides an overview of GIS technology, including GIS software programs (LITA, n.d.). ALCTS has also offered webinars such as, “Transitioning from cataloging to creating metadata” (ALCTS, “Transitioning,” n.d.) and “Using Open Refine to update, clean up, and link your metadata to the wider world” (ALCTS, “Using Open Refine,” n.d.).

Journals

Discipline-specific journals may contain articles that discuss digital humanities as related to a particular field. Journals focusing on digital humanities more broadly are also useful resources for metadata and catalog librarians new to the field of digital humanities, presenting useful case studies and information about developing new skills to support digital humanities projects.

*Digital Humanities Quarterly* (DHQ) is “an open-access, peer-reviewed, digital journal covering all aspects of digital media in the humanities” (Alliance of Digital Humanities Organizations, “About DHQ - Overview,” 2014). Published since 2007, *Digital Humanities Quarterly* is a forum for digital humanists from around the world, and provides for a wide range of submissions including articles, editorials, interactive media, and reviews of publications, conferences and other resources (Alliance of Digital Humanities Organizations, “DHQ submission guidelines,” 2014).

The *Journal of Digital Humanities* (JDH) is another open-access, peer-reviewed journal focused specifically on advances in the digital humanities. The articles published in the *Journal of Digital Humanities* originate in the online “experimental, edited publication,” *Digital Humanities Now* (“Digital Humanities Now - About,” n.d.).

Non-digital humanities journals occasionally have special issues devoted to topics in digital humanities. One example of this is the special “Digital Humanities in Libraries” issue of the *Journal of Library Administration* published in 2013, which included a variety of articles on ways that those working in libraries can get involved in digital humanities projects.

code4lib publishes Code4Lib Journal (C4LJ), an open access publication. Published four times a year, this journal includes a wide range of topics, including case studies and use cases, information regarding special metadata topics such as geospatial metadata, and methods of implementing metadata creation. Metadata and catalog librarians may also benefit from some of the non-metadata specific articles that may help acquaint them with specific database systems and other terms they may encounter such as SPARQL, Drupal, and broader uses for RDF.

Email Lists and Listservs

A great deal of useful discussion about digital humanities takes place over project- and topic-specific email lists. These lists fall into several areas including technical skills, general digital humanities topics, and library-oriented lists.
XML-DEV is a mailing list focusing on developing XML code and discussion of its use. Although it may be a bit technical, this list can be beneficial for metadata and catalog librarians because it is also used by people asking project-specific questions about using XML. Depending on their repository system, metadata and catalog librarians may benefit from signing up for the email lists dedicated to using it. Email lists exist for a variety of systems, including Fedora, DSpace and two email lists for Islandora (islandora and islandora-dev). centerNet sponsors a website dedicated to providing a list of resources for starting digital humanities programs (centerNet, n.d.). The DHCenterStartUp list and the centerNet list provide useful forums for discussion related to creating digital humanities programs.

Some library-specific email lists can provide a forum for metadata and catalog librarians to post questions related to metadata for digital humanities. These include metadataLibrarian and DIGLIB. Subscribing to metadata schema-focused email lists such as the MODS Listserv and DC [Dublin Core]-General Mailing List may be beneficial to learn more about implementing and adapting these metadata schemas for faculty projects. The ACRL Digital Humanities Discussion Group is a discussion forum geared specifically towards academic librarians involved in digital humanities, and can help provide metadata and catalog librarians with background information needed to communicate effectively with faculty and staff who are involved in digital humanities projects on campus.

BUILDING AND SUPPORTING DIGITAL HUMANITIES SYSTEMS ON CAMPUS

As interest in the digital humanities has grown rapidly among faculty and college administration, librarians and technologists on campuses around the country have begun to develop methods of supporting faculty projects in digital humanities. Because of this rapid growth, best practices and suggestions for how to support these types of projects, particularly on smaller college campuses, are still being developed. Unless new positions are created in the library specifically to support work in digital humanities, the task of assisting with these projects must be added to the workload of current librarians and library staff. Although this may seem daunting, skills and resources that already exist among librarians and technologists can be repurposed to support the technological infrastructure and research and project development needs of faculty projects.

Campus-wide agreements may need to be reached about the technological infrastructure designed to support digital humanities on campus. It is important that metadata and catalog librarians play a role in evaluating the database platforms and workflows that are being considered, particularly if all digital humanities projects are to share the same system. If a single method is going to be implemented across campus, it is important that many aspects be investigated to ensure that the selected products are flexible enough to accommodate a variety of projects, and to allow faculty and students to work with ease in the system. In addition to evaluating the types of metadata that can be utilized by a system, it is also important that metadata and catalog librarians examine the ease of metadata input. Points to examine include:

- How is metadata added to the database? Can records be easily batch loaded or are they added one at a time?
- Is there a user-friendly interface for adding/editing metadata, or are technological skills in XML or another encoding standard necessary for accomplishing this?
- How can the metadata be used for searching in the database? Are customized or faceted searches available that take advantage of robust, full metadata?
Beyond the Back Room

- Is new or edited metadata immediately published to the active site, or is it possible for it to be hidden from public view until it has been reviewed?
- How is the metadata lifecycle handled? Are metadata records available after an object is deleted?

Depending on the needs of the projects on campus, the answers to these questions may dictate the type of system that is selected. Decisions that are made in this regard will impact the abilities of the metadata and catalog librarian to work with faculty members in developing customized metadata schemas. The system that is selected can place limits on the metadata standard that is able to be used (MODS, DC, CCO, etc.) and this should be addressed before a system is officially adopted.

CONCLUSION

Metadata and catalog librarians have moved from cataloging analog objects, to cataloging digital objects using schemas such as Dublin Core, MODS, CCO, VRA Core and others. These cataloging and organizational skills possessed by metadata and catalog librarians are a valuable contribution to scholarship in digital humanities projects on campus. Although new skills in areas such as TEI and XML may be necessary to fully support faculty projects in digital humanities, a variety of resources and learning opportunities are available to assist in this process.

With their extensive and diverse knowledge of metadata schemas and controlled vocabularies, metadata and catalog librarians offer a wide variety of potential contributions to digital humanities projects. They can interface directly with faculty and students about information organization, how searching capabilities are influenced by metadata, training in metadata creation and good organizational practices for metadata. Although they are frequently hidden away in the “back room” of the library, metadata and catalog librarians have much to offer and make ideal partners in the growing field of digital humanities.

REFERENCES


The Digital Humanities initiative (DHi) at Hamilton College: Creating a collaboratory for digital humanities in a liberal arts setting. (2010). Unpublished Andrew W. Mellon Foundation grant application, Hamilton College, Clinton, NY.

Beyond the Back Room


**ADDITIONAL READING**


**KEY TERMS AND DEFINITIONS**

**Dublin Core:** A metadata schema that is frequently used to record basic descriptive metadata.

**Drupal:** An open source content management system that can be used to build websites through the integration of various modules.

**Fedora Commons Repository:** An open source repository that can be used to organize a wide variety of types of digital objects.

**Islandora:** An open source system that features various solution packs and is designed to serve as a middleware between Fedora and the front end display of a repository.

**MODS:** Metadata Object Description Schema; a schema for encoding metadata that was developed by the Library of Congress in 2002.

**OCR:** Optical character recognition; an automated process of creating a transcript of text found in an image.

**TEI:** Text Encoding Initiative; a method of encoding text that provides layers of searching and interpretation by using XML for markup.

**XML:** EXtensible Markup Language; a flexible method of encoding data using customizable tags that can be used in conjunction with controlled vocabularies and encoding standards.

**ENDNOTES**

1 To see the description of the MODS record template field-by-field, please visit: http://wiki.dhinitiative.org/doku.php/metadata/metadata1. Other supporting metadata documentation including links to controlled vocabularies and standardized file naming conventions can be found under the “Metadata” heading of the Digital Humanities Initiative Documentation Wiki homepage (http://wiki.dhinitiative.org/doku.php/).