

University of Massachusetts Amherst

From the Selected Works of Jennifer Eustis

2016

Transforming Roles for Catalog/Metadata Librarians through New Initiatives

Jennifer Eustis, *University of Massachusetts Amherst*



Available at: <https://works.bepress.com/jennifer-eustis/9/>

Transforming Roles for Catalog/Metadata Librarians through New Initiatives:
Research Data, Digital Humanities, and the Digital Repository
at the University of Connecticut Libraries

Jennifer Eustis

Introduction

Catalog/metadata units are ever changing. New directions are being taken with how materials are being added to databases. Batch loading has changed from being only books and now includes electronic resources such as streaming media. Workflows and processes include diverse formats and purchasing options such as owned, rented, or leased materials.

Catalog/metadata units have incorporated these new workflows and processes into the more traditional work of binding, labelling, and cataloging individual unique items. In addition to unique materials, these units also participate in managing metadata for digital collections. To add to this variety are the new content standard, Resource Description and Access (RDA),¹ and the bibliographic framework BIBFRAME.² Suffice it to say that catalog/metadata units are moving forward and transforming. While this diversification has allowed for new opportunities, new roles are there to be explored.

The realm of digital scholarship provides a number of these new roles for catalog/metadata librarians. At many institutions, these librarians have helped implement, create, or enhance records for digital collections. Though the work varies from one institution to the next, in general these units adapt cataloging skills to digital initiatives. These cataloging skills consist of organizing information, working with metadata standards, and creating consistent and accurate data for access and discovery. These projects have allowed catalog/metadata librarians to step beyond MARC21 and their institution's library management system. Thanks to the realm of digital scholarship, their expertise can be further extended beyond library-centric collections

and myopic services.³ Opportunities to work with researchers and discipline-specific data and tools are all present in providing support for the needs of digital scholars.

This chapter will highlight three digital scholarship initiatives at the University of Connecticut Libraries (UCL) that have radically changed the role of librarians in the Cataloging and Metadata Services Unit. The first initiative is the digital repository. Much more than an institutional repository, the Connecticut Digital Archive's (CTDA) primary focus is to provide digital preservation services to Connecticut-based non-profit cultural and memory organizations, including libraries, archives, galleries, and museums. The digital repository also provides digital preservation services for the University of Connecticut's (UConn) research data output. The second initiative is the Digital Humanities and Scholar's Collaborative program, which offers workshops, consultations and other services for those working in the digital humanities. The third initiative will illustrate the contributions of unit members to support research data. New roles played by unit members in UCL's Cataloging and Metadata Services (CMS) unit and the lessons learned will be discussed with each initiative.

The CTDA: A Digital Repository

In 2011, UCL expressed the desire to go beyond mere access of digital materials to their long-term preservation. At that time, UCL had outsourced solutions providing access to digital resources found in archive and special collections, as well as electronic theses and dissertations. UCL's outsourced solutions presented several challenges. These solutions could not be used for a wide variety of formats. While complex objects and sound recordings were attempted to be ingested, long upload times and frequent error messages made this process slow and unreliable. File size limitations of 2mb per file also presented several challenges. This was suitable for text and smaller images, which typically range from 500kb to 1.5mb for high-resolution images;

however, research data, streaming video, audio, and other digital formats are often larger than 2mb. Moreover, masters were not preserved which led to a decision to store and manage masters on a separate archival file master server. This was a poor solution, because it led to more work and trying to track objects in two separate environments. Finally, the work was typically funneled through one individual, creating a bottleneck at one point in the entire process.

In 2011 a working group was formed to recommend a new digital repository solution for UCL's growing digital collections and research data. This group endorsed and later implemented two open source solutions. The Fedora Digital Repository is an open source flexible and extensible architecture used to store, manage, and access digital content in any format. Working on top of this layer is Islandora, an open source framework used to manage, deliver, and present diverse digital content. This combination provides UCL with a powerfully flexible, extensible, and adaptable digital repository that has the possibility of the Trustworthy Repositories Audit & Certification: Criteria and Checklist (TRAC)⁴ certification. Implementing Fedora and Islandora required developing a foundation consisting of policies, documentation, workflows, and procedures in multiple areas and by multiple teams.

For metadata, this meant creating implementation guidelines, data entry forms, display format, and determining how metadata would be delivered and indexed. This required several steps. An initial decision was made about the types of standards used. For the Catalog and Metadata Services Unit (CMS), these standards directly concern metadata used to describe and harvest resources. It will also involve adding preservation metadata in the near future. It was necessary to take into account the need for metadata to be flexible and adaptable for data coming from various types of institutions or projects such as libraries, historical societies, state agencies, or research data. CMS decided early on to implement the Metadata Object Description Schema

(MODS)⁵ as the normalized descriptive standard to describe resources in CTDA. CMS chose to implement simple Dublin Core⁶ for the Open Archives Initiative Protocol for Metadata Harvesting (OAI).⁷ Finally, a review of technical metadata was performed in collaboration with other units to ensure consistency throughout the project.

After this initial decision, CMS' attention turned to create implementation local practices for the descriptive metadata. These guidelines provide both a metadata application profile and an explanation of how UCL use MODS in the CTDA. Moreover, they outline other content standard requirements and recommendations for fields including subjects, genres, and dates. The audience for this document was and continues to widen, because it now includes CTDA partners from state agencies, museums, historical societies, cultural centers, and libraries. This document is also aimed at those that harvest CTDA descriptive metadata such as the Digital Public Library of America or EBSCO.

Finally, configuring Islandora's data entry forms and metadata display was crucial. Out of the box, Islandora comes configured for Dublin Core. This meant creating data entry forms based on MODS. These forms also had to take into account the particular user group that would be using these forms. As a result, several forms were created to suit the needs of several audiences. Because the CTDA is a multisite installation of Islandora, this added a layer of complexity to the MODS data entry forms which needed to be created and maintained in a number of sites. Each site also has their preferred metadata display, or which MODS elements are displayed with each record. Just like the forms, the metadata display configuration requires not only the input of the users but also user testing.

From the start of the project, user input resulting from consultations have been at the crux of metadata work. This involves direct interaction with members of other library units, CTDA

partners, and various stakeholders on campus. Consultations cover a variety of purposes. They assess users' needs; sometimes they involve education and training. Other times, they require moving metadata from one system to the CTDA. While needs vary, users seek out metadata librarians for their expertise with handling data. Metadata work now offers services such as data migration, creation, or enhancement, education and training, and help with documentation.

The description above illustrates that librarians in CMS have had to take on new roles. The most significant one is moving from being a back room operation to a public service desk. Metadata work for the CTDA consists of direct interaction with users. These transactions can be with fellow staff, CTDA partners, or researchers. This work does not take place in a vacuum but relies on the input of these content owners and knowledge experts. The consultation acts as a mediation between users' needs and the limits and requirements of the repository. The other new role is being a data manipulator. Though not exactly new, librarians in CMS now frequently interact with data in formats other than MARC21 and encoded with library-specific standards. In this role, CMS librarians have adopted their skill set to new types of data. Finally, there is the role of systems configuration. This includes metadata display, OAI harvesting, and MODS data entry forms. These roles have allowed CMS to further diversify its skill set and services to the UConn community.

More importantly, these roles have been an opportunity to showcase the value of a technical services background. Catalog/metadata librarians are excellent at organizing information according to a variety of standards and local practices. They have worked for years on developing efficient procedures and workflows. They have adapted to new formats and environments over the years, even when faced with decreasing resources. These new roles bring CMS librarians out of the shadows. Fellow colleagues and other users can better see what

metadata work is and how that work can help in the repository. Consequently, these new roles help make the case for more resources.

The Digital Humanities & UCL's Scholar's Collaborative

In 2012, UCL created a working group called the Scholar's Collaborative to respond to faculty and students' needs in the digital humanities. The Scholar's Collaborative brought together a diverse group of library staff: IT, the Music and Dramatic Arts Library, and the Digital Curation Librarian. This team provided education and training and support for pilot projects. Education and training covered tools helpful to projects in the Digital Humanities such as Omeka,⁸ Neatline,⁹ or Annotation Studio.¹⁰ The success of this team led to the creation of the Scholar's Collaborative program and space, coordinated by the Music and Dramatic Arts librarian.

CMS librarians, though not directly involved, cooperated on many projects with the Scholar's Collaborative team. The first collaboration was a workshop series entitled "TEI in July." Many digital scholars at UConn were and continue to be faced with understanding the markup language the Text Encoding Initiative (TEI).¹¹ This first workshop brought together librarians and scholars to work on marking up transcripts from logbooks. These logbooks were part of a digital humanities project from one of the scholars attending the workshop. Though the workshop was led by a CMS librarian, it promoted a collaborative and informal learning environment. This workshop was very successful in that attendees gained a better knowledge of TEI and helped a PhD candidate publish her work in TEI online. The success of this workshop has led to offering this series every year in July.

The role played by CMS librarians for the Scholar's Collaborative is consultant. Serving as an advisor involves education, training, and being available for consultations. A CMS librarian often co-teaches Scholar's Collaborative workshops when workshop topics relate to metadata. When the workshop is entirely on metadata, such as the workshop on Open Refine,¹² a data cleaning and profiling tool, a CMS librarian leads it. As a resource on metadata, the coordinator of the Scholar's Collaborative often calls on a CMS librarian to get answers on questions on metadata or to ask that a librarian come to a consultation. CMS librarians are also sometimes called to provide advice on metadata questions for data management plans in the digital humanities.

This collaborative role has added depth and complexity to the changing work of CMS librarians. The CTDA providing the opportunity to develop new services and a new service model based on directly interacting with users. Being a partner to the Scholar's Collaborative program emphasizes the importance to model services to interact directly with patrons. Traditionally, decisions made on what metadata standards to implement was originated through a library-centric model. Now, the digital scholars are driving the agenda for the Scholar's Collaborative. Many have an idea of what metadata standards they want to use, such as TEI. The new model adopted by CMS forces librarians to think beyond just library-centric standards, procedures, policies, and workflows. Decisions on metadata are made in concert between digital scholars and metadata librarians. Each party provides guidance based on their knowledge and expertise.

Research Data at UCL

The question of how to support research data at UCL began in earnest with the announcement in 2011 from the National Science Foundation (NSF)¹³ on requiring data

management plans in grant applications. UCL created a team to be part of the Association of Research Libraries' (ARL) eScience Institute.¹⁴ The team included science subject specialists, a CMS librarian, a health science librarian, a member from UConn Sponsored Program Services, and a member of UConn's IT department. A strategic document resulted from this learning experience; its primary recommendation was the creation of an E-Science team consisting of a broad range of participants from throughout UConn. This team was created in 2012.

The charge of this team was education, training, consultation, and data management plan review. The first initiative was in education and training. One subject specialists in the library had already started providing workshops before and during the ARL E-Science Institute. Relying on that experience, the workshops were expanded and included the collaborative efforts of the eScience team. The first year focused on data managed plans. In the second year, these workshops were broadened to include topics on managing research data such as data security, storage, archiving and preservation, metadata, legal and ethical concerns, or the data life cycle. The workshops were successful and drew on average 20 to 80 for one workshop.

Because of workshop evaluations, the section on metadata evolved greatly. This transformation mirrors that of the CMS from a library-centric approach to a service model directed and driven by digital scholars. At first, the presentation on metadata was library-oriented. It introduced metadata records and Dublin Core as a possible standard to use to describe data sets. It also did not provide examples except for records found in ULC's digital collections or catalog. As more evaluations came in and more contacts were made, the presentation integrated more examples from the sciences such as actual data sets in the form of spreadsheets. Moreover, everyday examples from Amazon records to readme files were introduced as concepts for "recording" descriptive information about research data. Another

development concerned the word metadata itself. As the presentation evolved, metadata was supplanted with the terminology of contextualizing or providing a framework of understanding and reference for one's research data. These small steps were well received by participants. They also mirror the change in focus from library-centric terms and approaches to looking at metadata from other perspectives.

These transformations entailed adopting new roles in CMS. These roles are also present in the two other initiatives above, the CTDA and Scholar's Collaborative. Namely, CMS librarians take on the roles of collaborator, educator, trainer, advisor, and service provider. Furthermore, these new roles have encouraged knowledge cross pollination between staff and knowledge experts. CMS librarians have learned and continue to learn more about metadata thanks to these new roles and the new service model in place. CMS librarians are now center stage in the library thanks to their participation in these public events. With this change in focus, CMS librarians are better able to communicate their value to fellow staff and library administration.

New Roles

These digital initiatives have transformed the roles played by CMS librarians. One role is that of trainer and educator. This role is often played out in one of UCL's learning spaces. CMS librarians as educators engage not only their fellow colleagues, but also members of the UConn community to help provide guidance and advice on metadata. CMS librarians play the role of content herders, producers, and disseminators. Librarians now offer services of data management that include migration, creation, and enhancement. They help migrate (or herd) data from one system to another. These data are manipulated in various systems, which leads to another role played, i.e. system configuration. This latter pertains to putting finishing touches on systems to

help with discovery and access. Visualizations, consistent faceted searches, access and discovery come about thanks to the role that catalog/metadata librarians play in configuring and developing systems to present digital assets. This leads to another essential role, organizational planners. Having consistent and accurate metadata is half the story. The other half is knowing how to organize data or digital assets and where to organize them for maximum effect. This highlights why creating collaborative networks and partnerships is essential.

In a presentation to the Association for Library Collections & Technical Services (ALCTS) Development Committee in January 2003, Martin Kurth outlined the goals of publishing metadata services at Cornell that began as early as 2002.¹⁵ From the beginning of these services, consultations played center stage. In slide 3, Martin writes: “[...] Establish and operate a ‘consulting to production’ metadata service capable of producing metadata in a variety of formats to organize, manage, and preserve collections over time and to enable effective discovery and use” (Kurth, 2003).¹⁶ Even three years later, Kurth’s presentation on the lessons learned from launching these services emphasize the dominant role of consultations. In slide 6 he writes: “The Metadata Services provide metadata consulting, design, development, production and conversion services to Cornell’s faculty, staff, and community partners to increase the value of their digital resources” (Kurth, 2004).¹⁷ CMS librarians have learned as well that consultations are the foundation from which services spring.

Lessons Learned

Transforming the unit and taking on these new roles has not come without lessons learned. It could be argued that these roles are anything but new. For the sake of argument, this is indeed true. Catalog/metadata librarians have been and continue to act as advisors, consultants, educators, trainers, system builders, organizational planners, partners, and collaborators. What is

new, however, is not the roles per se but the environments within which these roles are played out. These new environments belong to the realm of digital scholarship. They are environments in which you find digital scholars. This transformation entails several key steps: revamping outdated perceptions of catalog/metadata librarians and developing a sustainable service model. Too often, catalog/metadata librarians are pigeonholed into a negative stereotype. This latter often portrays these librarians as stodgy, rule-bound enthusiasts bent on applying library cataloging rules to the letter. Frankly, the damage of stereotypes is not the fictions they weave, but the fact that many are apt to believe them or parts of these tall tales. It is necessary to find out what tales are being told at your institution. CMS librarians found that they had been branded a back-room operation and were intrinsically not a part of any public service. Collaborating with fellow colleagues, co-teaching workshops, providing workshops, and offering services helped reshape another story of CMS. This restructuring happened over the course of approximately two years. In that time, a positive message of collaboration, partnership, and a dedication to all users, staff and community alike, was and is communicated repeatedly.

Changing others' perspectives of CMS also meant transforming the unit's image of itself. Mitigated fear was a first step. Like many institutions, it was unclear how to balance learning new skills, offering new services, and managing the other day to day processes. CMS started with a small project and learned a skill specific to that project. With each new project, new skills and more experience was gained. The CMS unit also mitigated its own fears of change by providing support for each team member. The unit realized the need also to become more flexible in terms of being able to take on a variety of different tasks and interaction with new and diverse users. But this flexibility and evolution demanded that an investigation on what could be stopped or done differently. These are extremely hard questions and often uncomfortable;

however, they are necessary to move forward. With these risks come several opportunities that go beyond new roles, namely a boost in confidence and an outward collaborative focal point.

Though CMS had services, they were library-centric. Services were primarily oriented for UCL's integrated library system. For those digital collections where CMS created and enhanced metadata, this was always done at the behest of another colleague in the library. The driving force was ILS projects. The driving standards were those created by cataloging librarianship. The focus was internal. With these new services of consultation, data management, education, training, and best practices, the driving force is the user, whether it be a fellow staff person or a digital scholar. The driving standards are those developed in the digital scholars' disciplines. The focus here is external. CMS librarians have to meet scholars where they are in their own environments for these services to work. That is why consultations act as a keystone, because it is essential to understand these environments to a certain extent to be able to advise on metadata in general. Metadata is not one-dimensional or library-centric, as Jenn Riley and Devin Becker so expertly illustrated in their "Seeing Standards: A Visualization of the Metadata Universe" (Riley and Devin, 2010).¹⁸ The world of metadata is complex and multi-dimensional, because one size does not fit all.

The other greatest transformation was to adopt a fail-fast fail-often practice. This convention meant making mistakes and learning from those mistakes to advance. It is also bound to the idea of being flexible. The saying that the customer is always right has a lot of truth contained in it. CMS' customers worked in their own environments, with their own habits, policies, procedures, and methods. Much like visiting a foreign country, working in or alongside these new environments required respecting them even if CMS librarians did not understand them or did not agree with them. Further, it entailed ensuring the sustainability of these services.

No one plan or service is full proof. The realm of digital scholarship is changing. Therefore, metadata services are always evolving as well. As much of this transformation is driven by technology, change happens more frequently. This is something that CMS librarians needed to understand and now deal with.

Conclusion

The transformation of CMS has not been without its challenges. CMS librarians have taken the time to learn new skills. They have also on occasion spent more hours on these new projects. These challenges and the lessons learned do not outweigh the results. CMS librarians have new roles and revamped skills that put them in center stage. Fellow colleagues see CMS librarians as full partners in fulfilling the libraries' goals. The perspective and old stereotypes have been replaced. CMS is now seen as a unit that provides services directly and indirectly to users wherever these users may be. These transformations have come about because of hard work. More importantly, these changes are a direct result of these new digital initiatives at UCL. It has taken CMS librarians almost four years to be able to construct a metadata program. It has also taken that amount of time to create a network of fellow staff and digital scholars who rely on and appreciate CMS librarians' expertise. All of this began with only one initiative, namely helping with research data workshops on metadata. It is not the size or amount of projects that count, but the effort that allows a department to transform itself and grow. CMS' new roles and new business model have made the unit more dynamic. They have even helped change old stereotypes and negative perceptions. More importantly, they have allowed catalog and metadata work, both the new and old, to become visible in a way that has not been the case for many years that benefit both UCL and UConn as a whole.

¹ Library of Congress, “Resource Description and Access (RDA): Information and Resources in Preparation for RDA”, <http://www.loc.gov/aba/rda/>, accessed November 19, 2014.

² Library of Congress, “Bibliographic Framework Initiative: BIBFRAME”, accessed November 19, 2014, <http://www.loc.gov/bibframe/>.

³ Bradford L. Eden, “From Our Readers: The New User Environment: The End of Technical Services?”, *Information Technology and Libraries* 29:2 (2010): 94.

⁴ Center for Research Libraries, “TRAC and TDR Checklists”, accessed November 19, 2014, <http://www.crl.edu/archiving-preservation/digital-archives/metrics-assessing-and-certifying-0>.

⁵ Library of Congress, “MODS: Metadata Object Description Schema”, accessed November 19, 2014, <http://www.loc.gov/standards/mods/>.

⁶ Dublin Core Metadata Initiative, “Dublin Core Metadata Element Set, Version 1.1”, accessed November 19, 2014, <http://dublincore.org/documents/dces/>.

⁷ OAI Executive, “The Open Archives Initiative Protocol for Metadata Harvesting”, accessed November 19, 2014, <http://www.openarchives.org/OAI/openarchivesprotocol.html>.

⁸ Roy Rosenzweig Center for History and New Media, George Mason University, “Omeka”, accessed November 21, 2014, <http://omeka.org/>.

⁹ Scholars’ Lab, University of Virginia Library, “Neatline: Plot your Course in Space & Time”, accessed November 21, 2014, <http://neatline.org/>.

¹⁰ MIT HyperStudio, “Annotation Studio”, accessed November 21, 2014, <http://www.annotationstudio.org/>.

¹¹ TEI Consortium, “TEI: Text Encoding Initiative”, accessed November 21, 2014, <http://www.tei-c.org/index.xml>.

¹² “Open Refine”, accessed November 21, 2014, <http://openrefine.org/>.

¹³ National Science Foundation, “Chapter II – Proposal Preparation Instructions”, accessed November 21, 2014, http://www.nsf.gov/pubs/policydocs/pappguide/nsf11001/gpg_2.jsp.

¹⁴ Association of Research Libraries, “E-Science Institute”, accessed November 21, 2014, <http://www.arl.org/focus-areas/e-research/e-science-institute#.VG-ptovF98E>.

¹⁵ Martin Kurth, 2003, “Establishing a Metadata Service”, <http://hdl.handle.net/1813/2663>, Presentation to the Library Collections and Technical Services Leadership Development Committee, January 2003, Philadelphia, PA, accessed August 18, 2014.

¹⁶ Martin Kurth, “Establishing a Metadata Service”, slide 3.

¹⁷ Martin Kurth, 2004, “Three Years Later: Lessons Learned from Establishing a Metadata Service”, accessed August 18, 2014, <http://hdl.handle.net/1813/1479>, slide 6.

¹⁸ Jenn Riley and Devin Becker, 2010, <http://www.dlib.indiana.edu/~jenlrile/metadatamap/>, accessed November 22, 2014.