

A Clean Sweep: The Tools and Processes of a Successful Metadata Migration

Anna Neatrou, University of Utah, J. Willard Marriott Library, anna.neatour@utah.edu

Jeremy Myntti, University of Utah, J. Willard Marriott Library, jeremy.myntti@utah.edu

Matt Brunsvik, University of Utah, J. Willard Marriott Library, matt.brunsvik@utah.edu

Harish Maringanti, University of Utah, J. Willard Marriott Library, harish.maringanti@utah.edu

Brian McBride, University of Utah, J. Willard Marriott Library, brian.mcbride@utah.edu

Alan Witkowski, University of Utah, J. Willard Marriott Library, alan.witkowski@utah.edu

Abstract

In 2016, the University of Utah's J. Willard Marriott Library migrated digital asset management systems from CONTENTdm, a vendor provided solution from OCLC, to Solphal, a homegrown system utilizing several open source tools. During the migration, issues with metadata led to a large-scale metadata cleanup and standardization project, enhancing discovery in our new system. This article discusses the method used to determine which system would best meet our needs, methods for metadata migration, issues observed during migration, metadata management capabilities of the new system, and future plans for post-migration metadata cleanup and remediation to ensure that our metadata is consistent with best practices.

Keywords

metadata; DAMS; systems migration; open source software

Introduction

Systems migrations are an inevitable necessity over time when needs and technology change, as libraries continuously strive to improve the discoverability of their digital collections. The University of Utah's J. Willard Marriott Library was an early adopter of CONTENTdm as its digital asset management system in the early 2000s. Due to changing needs and scalability issues, it was decided that a migration away from CONTENTdm was necessary in order to continue providing the best possible access to our digital assets. The new system that has been developed incorporates open source software using Apache Solr (indexer), NGINX (web server), and phalcon (PHP framework). During the course of migration in 2016, several issues with metadata were exposed from many legacy collections created over the past two decades. To deal with these metadata issues, a large scale metadata cleanup and standardization project was completed during the migration to make the data was more accurate and user friendly. The new system for our digital library officially went live in December 2016 and can be viewed at <https://collections.lib.utah.edu/>. This article will discuss the method used to determine which system would best meet our needs, how the data was migrated along with metadata issues observed during migration, the metadata management capabilities of the new system, and future plans for post-migration metadata cleanup and remediation to ensure that our metadata is consistent with best practices.

Literature Review

Migrating from one digital asset management system (DAMS) to another requires a considerable amount of investment in time and effort on the part of library staff. In Taking Control: Identifying Motivations for Migrating Library Digital Asset Management Systems, Stein and

Thompson explored reasons for migration by reviewing case studies of migrations and developed a survey designed to capture reasons for migration, showing a general trend for migration to open source systems that can support local control of customizations and extensions to DAMs functionality (Stein and Thompson 2015b).

CONTENTdm has long been a popular DAMS option for digital libraries, but limitations and the costs involved in licensing vendor-based options cause digital library managers to look for other options. The Lowcountry Digital Library undertook a migration away from CONTENTdm when faced with additional costs associated for a CONTENTdm collection that grew past the limit of 50,000 items (Gilbert and Mobley 2013). The need for greater flexibility and standardization contributed to Simon Fraser University Library migrating to Islandora (Jordan 2015). Oregon Digital's previous use of CONTENTdm required local modifications to support search and discovery and functionality for users, and the desire for more flexibility resulted in a project to migrate their system to Hydra. This migration also incorporated the development of a local authorities source, OpaqueNamespace (Simic and Seymore 2016). A DAMS review process at the University of Houston identified Hydra/Fedora as a preferred solution, with a phased migration from CONTENTdm involving system installation, metadata migration, and interface development (Wu et al. 2016). The Robert R. Woodruff Library developed a matrix for ranking DAMS according to features such as presentation, preservation, and repository, weighing these alongside cost and time involved in implementing open source, vendor-based, and hybrid options. Their proposed digital library solution includes support for ArchivesSpace, usage of XTF for EADs, usage of Fedora 4 and Duracloud for preservation, and CONTENTdm and Omeka as presentation layers (Wiseman and Matthews 2016). This tendency to maintain a

variety of solutions to fully support all the desired functionality of a digital library is also shown in a project at Texas A&M, which concluded that rather than a single DAMs, a Digital Asset Management Ecosystem was required (Bailey et al. 2016). Stein and Thompson found the main metadata needs for a future DAMS were threefold: supporting a variety of metadata schema, support for exporting and metadata reuse for library staff and users, and the need to integrate digital object identifiers (Stein and Thompson 2015a).

Metadata assessment and remediation are important aspects of a DAMS migration. Often the literature points to a need for developing a comprehensive framework for metadata quality, supported by tools and machine processing methods that can be used for assessment (Tani, Candela, and Castelli 2013). Recently the DLF Metadata Assessment Group has completed an environmental scan covering both the literature as well as tools and software packages used in metadata assessment. Common themes found as part of this environmental scan showed people engaged in developing frameworks, developing methods of enriching existing digital collections, capturing changes to metadata, measuring quality of auditing, and defining considerations for metadata quality in an aggregated or shared environment (“Digital Library Federation (DLF) Assessment Interest Group (AIG) Metadata Working Group” 2016). Work at the University of Houston demonstrated innovative ways of leveraging the CONTENTdm API in preparing for metadata migration (Wu et al. 2016). Metadata aspects of a DAMS migration at Yale focused on what was practically possible, if it was easy to fix metadata using bulk transformations before migrating it was completed, but other fixes were planned for post-migration (Dula 2014). When Duke University Library moved to Hydra from a previous open source system using Django, metadata remediation was accomplished with a variety of techniques. Using tools like Google

sheets to capture comments about metadata fields, Tableau for analysis, combined with regular expressions, and scripting with Ruby, metadata enhancements like implementing EDTF were completed for their migration (Dickson, 2016). While aspects of requirements criteria for DAMS, reasons for migrating, and frameworks for metadata assessment often share points in common across different projects and institutions, methodologies and technology frameworks are necessarily customized to fit the needs of a local repository system. The University of Utah Marriott Library's migration to Solphal provides another case study illustrating the decisions and requirements for DAMS migration, as well as the steps taken in migrating a sizeable amount of legacy metadata.

Digital Library Overview

The scope of the Marriott Library's digital library migration was formidable, including 294 separate digital collections with over 787,000 individual items. Of these items, 133,289 were compound objects containing two or more files tied together as single objects housing thousands of more images. These items and images comprised a wide variety of 38 differing file formats. Together, they constituted 2.15 terabytes (TB) of online display storage. More daunting was the process of reaching a consensus on standardization among over 50 different hosting partners contributing to the digital library. Many colleges and departments within the University of Utah as well as institutions external to the University, such as the Utah State Department of Heritage and Arts and the Murray City Museum, utilize our DAMS, all with specific needs with regards to display, metadata, and discovery. We quickly realized that the scope of this endeavor would require the talents and expertise of multiple departments within the Marriott Library as well as collection manager and end-user feedback.

The process for coordinating between the various stakeholders of the migration followed a multifaceted communication effort. The first effort consisted of established email lists to inform and remind stakeholders of important events. The second effort consisted of creating a public webpage where stakeholders could access at their convenience in order to keep track of progress and important developments made during the migration process. The third effort consisted of in-person meetings and trainings to ensure everybody was on the same page, as well as had their voices heard with respect to the priorities of their unique collection needs. All three efforts are still in place as we continue to receive feedback on the new system.

SIMP Tool for Metadata Management

Upon beginning the process of implementing Rosetta for preserving digital assets in the fall of 2012, we discovered that a tool was needed to be able to streamline workflows for ingesting digital content in our DAMS and digital preservation system (DPS) along with the corresponding metadata. To tackle this issue, the Submission Information Metadata Packaging (SIMP) Tool was developed. The SIMP Tool was designed to be platform agnostic and modular since it was inevitable that a systems migration away from our current DAMS and DPS would happen in the future. Rather than focus exclusively on the systems currently being used which were required to connect to the SIMP Tool, the needs of the organization and users were taken into account so future systems migrations would still be able to utilize the functionality built into the tool. For more information on how the SIMP Tool was developed and used in 2014, refer to this previously published article (Neatrou et al. 2014). While many enhancements have been implemented in the SIMP Tool since this previous article was written, it gives a basic overview

of how the SIMP Tool works for preparing digital objects for ingestion into both a DAMS and DPS.

The SIMP Tool allows us to ingest intellectual entities (IEs), create descriptive and technical metadata (through manual and automated means), assign Archival Resource Keys (ARKs) as persistent identifiers, send the IEs to our DAMS as well as our DPS. When the SIMP Tool was first created and the library was using CONTENTdm as its DAMS, the workflow to ingest content into the DAMS required several steps to extract the metadata from the SIMP Tool and then process the files using the CONTENTdm Project Client including approval and data indexing. Even before the migration away from CONTENTdm, the SIMP Tool was able to replace several tasks typically completed in the CONTENTdm Project Client for metadata creation. With the migration to the new system, content is able to be ingested in Solphal by hitting one button which sends the IEs to the DAMS and indexes the new content nearly instantaneously.

The SIMP Tool has been instrumental in automating multiple metadata tasks. Several pieces of technical and preservation metadata are automatically extracted from the files, such as the file format, OCR text, and a checksum. Another recent metadata enhancement to the SIMP Tool includes the creation of controlled vocabularies. The first metadata fields that have been assigned controlled vocabularies within the tool are the Type field using the DCMI Type Vocabulary¹ and Format field using the Internet Media Type Vocabulary². Additional fields will have controlled

¹ <http://dublincore.org/documents/2000/07/11/dcmi-type-vocabulary/>

² <http://www.iana.org/assignments/media-types/media-types.xhtml>

vocabularies implemented in the future, such as the language field using the ISO 639-2 codes for the representation of names of languages³.

Within the SIMP Tool metadata editor, it is possible to edit the metadata for several IEs at the same time using a spreadsheet editor. The style of editor streamlines the process for creating new metadata as well as repurposing existing metadata from other sources such as Encoded Archival Description (EAD) files or other inventories using XSLT (Extensible Stylesheet Language Transformations). An additional enhancement to the SIMP Tool metadata editor includes validating data such as the date field to make sure it conforms to the ISO 8601 date and time format⁴. This includes a macro built into the tool to expand date ranges into multiple values (e.g. "1950-1955" becomes "1950; 1951; 1952; 1953; 1954; 1955").

Digital Asset Management System Review

After using CONTENTdm for more than a decade, there was a growing awareness that the size of our digital collections as well as our current practices with self-hosted CONTENTdm would not scale well into the future. Often urgent feature requests such as improved OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting⁵) support were not addressed, creating harvesting problems for sharing collections with the Mountain West Digital Library (MWDL) and the Digital Public Library of America (DPLA). Many of the digital library partners who host collections with our CONTENTdm repository had local customizations in their collections, which needed to be ported over with each software upgrade. Another critical architecture concern was the scalability issue related to extremely slow indexing services and the system requirements. Due to the large size of our collections and the indexing limitations of

³ <https://www.iso.org/standard/4767.html>

⁴ <https://www.iso.org/iso-8601-date-and-time-format.html>

⁵ <https://www.openarchives.org/pmh/>

CONTENTdm, it would take upwards of eight weeks to perform a full site re-index. Other issues that were of concern included the inconsistent user experience with loading pdfs, slow page load times, and the removal of features that were popular with patrons.

A committee was charged by the Marriott Library's Technology Services Council to provide an analysis of DAMS options, including CONTENTdm. In order to accomplish this, both internal and external stakeholders for our digital collections were consulted. The committee had members from a variety of areas in the library, incorporating feedback from special collections, public services, metadata, and digital operations. The work progressed through stages addressing stakeholder analysis, selecting digital repository software to be reviewed, defining requirements criteria, developing questions to ask vendors or assess against open source software feature sets, and creating a scoring model to use when ranking DAMS solutions (Masood and Neatrou 2014).

The process involved a variety of software options, some of which were true digital library repositories, and some which had features or elements the committee wished to consider in depth. The software and platforms reviewed, both full and partial possible solutions, included CONTENTdm (hosted and self hosted)⁶, Rosetta⁷, Equella⁸, Omeka (for digital exhibits)⁹, Cumulus¹⁰, ChronAm¹¹, Bepress¹², XTF¹³, Hydra¹⁴, MDID¹⁵, and Northwest Digital Archives,

⁶ <http://www.oclc.org/en/contentdm.html>

⁷ <http://www.exlibrisgroup.com/category/RosettaOverview>

⁸ <http://www.equella.com/>

⁹ <https://omeka.org/>

¹⁰ <https://www.canto.com/cumulus/>

¹¹ <https://github.com/LibraryOfCongress/chronam>

¹² <https://www.bepress.com/>

¹³ <http://xtf.cdlib.org/>

now Archives West¹⁶ for EAD files. The major considerations for reviewing the DAMS options are outlined in the following table.

[INSERT TABLE 1]

As we began evaluating the different Hydra heads and the underlying stack that Hydra is built upon, we developed a much better understanding of the core components. Beyond the investigation of the Hydra technology stack, we quickly became aware of the very large and vibrant Hydra community that was passionate about the work and helping introduce the technology to others. As we continued to forge ahead with the preliminary scoping of the project, it became very clear that SIMP will become the focal point for workflow related work. We began to consider the potential of developing a homegrown system that could leverage SIMP functionality, was based on a component from the Hydra stack, Apache Solr, and used familiar PHP frameworks to construct a new system. As we continued the evaluation of the homegrown system, it became apparent that given our tight timeframe, in house expertise, and other considerations, it would be best to develop a homegrown solution. We felt very comfortable that we could create the DAM, while incorporating our workflow management tool, the SIMP Tool, and replicate and expand on many of the functionality that our stakeholders valued in CONTENTdm.

Metadata Migration

¹⁴ <https://projecthydra.org/>

¹⁵ <https://sourceforge.net/projects/mdid/>

¹⁶ <http://archiveswest.orbiscascade.org/>

The major work for the migration of all of our digital library collections began in the late Spring of 2016 and was completed in December 2016. One reason to complete the migration in such a short amount of time was the need to fully migrate to the new system before server warranties expired on our CONTENTdm servers. The speed of the migration was also supported by the SIMP Tool, because there was no change for collection managers in the system used to manage collections and new descriptive metadata.

While a standard best practice in systems migration is to do the majority of metadata cleanup prior to migration, this short time frame to complete the migration did not allow for a detailed analysis of all of the metadata in order to make it as clean as possible. The majority of metadata cleanup and standardization completed prior to migrating consisted of standardizing metadata templates, normalizing field names, standardizing Dublin Core mappings, discarding unneeded or empty fields, and merging fields with similar data.

CONTENTdm provided the ability for collection managers to create custom metadata field templates for each collection. This flexibility ended up creating migration issues centered on field standardization. Before we embarked on the migration, our repository had 735 unique fields, many of which had local field names such as “subject 1”, “subject 2”, or “hidden description”. Many of the fields were not even used and contained no metadata values. Other issues centered on legacy metadata stored in collections that had not been updated or used in 10 years. Some of these collections had specific purposes at one time but were no longer needed such as limited access art history slide collections for specific classes, a book cover collection for

the local Career Services library, and the University Press catalog. By removing collections that were no longer needed, many of the bad metadata field labels were removed from the repository.

Metadata remediation followed a phased process that began with reviewing the collection templates. Generally as a local practice, new collections were created by repurposing a central standardized field properties template. However, individual collection managers were still able to edit and modify the templates, which led to inconsistencies. Collection managers created multiple and/or differing title, subject, and description fields when customizing their own templates. Some fields were misspelled. Some fields were typed as upper case, others as lower case, and contained extra characters or words. In order to migrate the data efficiently as well as for ideal faceting for enhanced discovery, the field naming needed to be consistent and standardized. The main purpose for this standardization was tied to search requirements. For example, in order to search title values across every collection, all collections had to contain a consistently named title field.

[INSERT FIGURE 1]

From this process, we established a protocol that relied on one central template of core fields to be used across all collections for faceting and queries. These core fields were based on the Mountain West Digital Library (MWDL) Dublin Core Application Profile¹⁷, such as title, creator, date, subject, rights, coverage-spatial, type, format, and ARK. In order to achieve this, the collection templates within CONTENTdm would have to be adjusted to reflect this protocol.

¹⁷ http://mwdl.org/docs/MWDL_DC_Profile_Version_2.0.pdf

However, we couldn't do this without first correcting further inconsistencies found in the metadata field labels and values.

The first step in correcting inconsistencies involved exporting inconsistent field name data to a column in Google Sheets. We then sorted data by collection alias to determine which collections needed naming changes in order to adhere to the standard we established. From that point, we only needed to edit according to standard in a separate column, and replace the affected metadata within CONTENTdm. Many of these changes could be completed through scripts that were run on the descriptive metadata files on the CONTENTdm server. Other changes that weren't straightforward had to be completed manually in the CONTENTdm administration interface.

[INSERT FIGURE 2]

The next step of field remediation involved identifying field labels that were commonly interpreted incorrectly. Many collection managers were confused as to what a proper value was for certain local field labels, which led to a host of fixes we had to perform. For example, we found description field labels whose values were actually subjects, and type field labels whose values were actually formats. For this, we reviewed our established core fields and relocated metadata values to the correct field as necessary. This included merging duplicate fields with similar metadata values into one core field, and deleting extraneous fields not used in collections across the repository. Once the specific fields in different collections were identified that needed to be merged, the task was completed programmatically directly on the server rather than needing manual work done where human error could have been introduced.

Our final step involved editing metadata values not in compliance with the MWDL Dublin Core Application Profile. For example, the type and format fields require specifically defined values from controlled vocabularies. However, there were multiple variations in syntax of these fields, which did not follow the correct standard. We sorted all noncompliant values on a spreadsheet and redefined them to the correct value in a separate column. Finally, the compliant values were entered back into the collections.

[INSERT FIGURE 3]

At this point, the collection templates within CONTENTdm were adjusted to reflect the basic standards of the new system. After all fields were cleaned up, there were 441 fields left, nearly 300 less than before the cleanup work was performed. Once the data was prepped and ready in CONTENTdm the data was consistent enough to be migrated to Solr. Using the new normalized field labels, the descriptive metadata in CONTENTdm was converted to valid xml and imported into Solr.

[INSERT FIGURE 4]

Prior to migration, some limited cleanup, assessment, and enhancement had been done on a collection by collection basis, but this was very limited in scale and the bulk of the metadata in the repository was migrated without an in-depth review of the metadata values. Future plans for metadata assessment and improvement include assessing required fields against the MWDL

application profile and developing targeted enhancements where possible, especially in adding additional values for fields required for harvest by MWDL and DPLA. In addition, the Marriott Library recently officially adopted Rightsstatement.org for digital collections¹⁸. While many new collections have rightsstatement.org statements in them, legacy collections have yet to be evaluated and copyright considerations for those items have not been researched, but improving discovery for digital library items based on rights statement properties is an additional future goal. Further post-migration metadata creation, clean-up, and enhancement work is facilitated by a variety of tools developed for working with metadata in our system Solphal.

Metadata management in Solphal

Metadata creation for all new content ingested into Solphal is completed through the SIMP Tool, providing consistency for metadata staff, as this was also the established method of creating new descriptive metadata in the previous DAMS. In order to edit metadata for items already ingested into Solphal, three tools have been developed to make metadata changes: a frontend interface for editing individual items, a mass update tool for making global changes based on different parameters, and methods for updating metadata directly in the Solr index.

Rather than having a separate backend interface for editing metadata for individual digital objects already ingested into the system, metadata can be changed directly from the frontend. If an authorized user is logged in to the system, an edit button will appear on each page which enables most metadata fields to be edited. Once saved, the updates are instantly made live and logged to a separate database for later review. The editing capabilities on the frontend also allow for deleting items, replacing the digital objects with new files, and recreating thumbnail images.

¹⁸ <http://www.lib.utah.edu/services/digital-library/policydigitalassets.php>

[INSERT FIGURE 5]

The mass update tool serves two purposes: easily determining duplicate or problematic metadata and the actual searching/replacing of metadata. The tool shows unique values and counts for each field, sorted either alphabetically or by the number of occurrences. Next to each value is an edit link, which allows for a global search and replace for that value in the specific field. The tool is ideal for fixing inconsistent or small variations in the metadata. This mass update tool can be used to do a global update of metadata across all digital collections, target individual collections for updating, or update data based on a specific query. This tool has already shown to be useful in post-migration metadata fixes for fields like type, format, and language, which rely on external vocabularies. Non-conformant values were quickly spotted, and these values were updated shortly after the tool was developed.

Metadata from searches can also be exported from the frontend to a tab-delimited text file. Edits can be made in a spreadsheet editor, then later imported directly into Solr using bash scripts. Currently we have used the tool to add subject headings for thousands of items in a large collection that required enhancement. We are also currently developing workflows for improved collection level faceting, by adding new FAST subject headings to selected collections based on existing Library of Congress Subject Headings values. Additional metadata cleanup issues for the future center around known duplicate images and non-distinct titles.

Conclusion

The overall success of the migration from CONTENTdm to Solphal reaped many rewards for both our collection managers and users. Collection managers are able to streamline their workflows to be more efficient, including several automated metadata tasks to help reduce human error as well as multiple options for editing metadata before and after ingestion. Through the use of the SIMP Tool, collection managers now have an easy to use tool for managing the metadata associated with their digital projects that is both scalable and platform agnostic, making it easier to adapt the tool for their future needs and additional systems. Users of the new DAMS have benefited from a site that is easier to navigate, has a much faster index response time, and has very limited downtime where digital collections are not available. The standardization of facets and metadata remediation through the migration has improved the consistency of fields across the digital library, enhancing discovery for digital library users.

As with any migration, there are challenges presented along the way. Since the new DAMS is using all open source systems and homegrown tools, the issue of continuous internal support and development will need to be addressed for the long term. With more people using the new system, more requests are submitted for further enhancement and technical support. Likewise, the more the SIMP Tool is used, the more we have seen bandwidth limitations regarding heavy use, such as with ingesting large audio or video files which take a long time for the tool to process. The articulation of these challenges is beneficial as long as we allocate the necessary resources to address them. The new DAMS allows for exponential growth, so an organized structure for sustaining this growth in both personnel and hardware needs to be addressed.

The migration's success was due to the cooperation and support of many people including members of the Digital Infrastructure Development and Digital Library Services Departments, as well as many internal and external partners of our library. By being able to quickly adapt to changes and new metadata issues discovered during migration, we were able to successfully complete the migration in a relatively short timeframe. By developing the system internally rather than relying on a vendor product, we are able to continue the development of the system to meet the needs of our collection managers, metadata catalogers, and users.

Bibliography

- Dickson, Maggie. 2016. "Looking Back, Moving Forward: Remediating Duke Digital Collections Metadata." presented at the ALCTS Virtual Preconference.
http://downloads.alcts.ala.org/mw_ac/ac16vp2_06082016_Automating_Legacy_Data_Cleanup_Projects_Dickson_Slides.pdf.
- "Digital Library Federation (DLF) Assessment Interest Group (AIG) Metadata Working Group." 2016. <http://dlfmetadataassessment.github.io>.
- Dula, Michael. 2014. "Digital Repository Development at Yale University Library." presented at the Coalition for Networked Information (CNI) Fall Meeting, Washington, DC, December. http://web.library.yale.edu/sites/default/files/files/cni_dec8_2014_yale3.pdf.
- Gilbert, Heather, and Tyler Mobley. 2013. "Breaking Up With CONTENTdm: Why and How One Institution Took the Leap to Open Source." *Code4Lib Journal* 20 (April).
<http://journal.code4lib.org/articles/8327>.
- Jordan, Mark. 2015. "SFU Library's Migration to Islandora." presented at the CNI Spring Meeting, Seattle, April 14. https://www.cni.org/wp-content/uploads/2015/04/Tues_Jordan_Innovative_Uses_Islandora.pptx.
- Masood, Kinza, and Anna Neatrou. 2014. "Digital Asset Management System Options: Report of the University of Utah Libraries DAM Review Task Force." Webinar presented at the Mountain West Digital Library Webinar Series, February 4.
http://mwdl.org/events/DAMS_options.php.
- Neatrou, Anna, Matt Brunsvik, Sean Buckner, Brian McBride, and Jeremy Myntti. 2014. "The SIMP Tool: Facilitating Digital Library, Metadata, and Preservation Workflow at the

- University of Utah's J. Willard Marriott Library." *D-Lib Magazine* 20 (7/8).
doi:doi:10.1045/july2014-neatrou.
- Simic, Julia, and Sarah Seymore. 2016. "From Silos to Opaquenamespace: Oregon Digital's Migration to Linked Open Data in Hydra." *Art Documentation: Journal of the Art Libraries Society of North America* 35 (2): 305–20. doi:10.1086/688730.
- Stein, Ayla, and Santi Thompson. 2015a. "Understanding Metadata Needs When Migrating DAMS." In , 119–28. <http://dcpapers.dublincore.org/pubs/article/view/3767>.
- . 2015b. "Taking Control: Identifying Motivations for Migrating Library Digital Asset Management Systems." *D-Lib Magazine* 21 (9/10). doi:10.1045/september2015-stein.
- Tani, Alice, Leonardo Candela, and Donatella Castelli. 2013. "Dealing with Metadata Quality: The Legacy of Digital Library Efforts." *Information Processing & Management* 49 (6): 1194–1205. doi:10.1016/j.ipm.2013.05.003.
- Wiseman, Christine, and Al Matthews. 2016. "Time, Money, and Effort: A Practical Approach to Digital Content Management." *AUC Robert W. Woodruff Library Staff Publications*. <http://digitalcommons.auctr.edu/libpubs/8>.
- Wu, Annie, Santi Thompson, Rachel Vacek, Sean Watkins, and Andrew Weidner. 2016. "Hitting the Road Towards a Greater Digital Destination: Evaluating and Testing DAMS at University of Houston Libraries." *Information Technology and Libraries (Online)* 35 (2): 5–18.