A Metadata Manager's Role in Collaborative Projects: The Rutgers University Libraries Experience

Rutgers University has made this article freely available. Please share how this access benefits you. Your story matters. [https://rucore.libraries.rutgers.edu/rutgers-lib/33943/story/]

This work is a SUBMITTED MANUSCRIPT UNDER REVIEW (SMUR)

This is the author's manuscript for a work which, at the time of deposit to RUcore, was under formal review managed by a socially recognized publishing entity. The manuscript may or may not have been subsequently published. Content and layout follow publisher's submission requirements.


Terms of Use: Copyright for scholarly resources published in RUcore is retained by the copyright holder. By virtue of its appearance in this open access medium, you are free to use this resource, with proper attribution, in educational and other non-commercial settings. Other uses, such as reproduction or republication, may require the permission of the copyright holder.

Article begins on next page
Abstract

Purpose – This article discusses the roles and responsibilities of a metadata manager in collaborative digital projects.

Methodology – It describes the general requirements for metadata management, and introduces some scenarios in the practices of digital projects by the Rutgers University Libraries to support the generalized definition. A workflow of metadata management is illustrated.

Practical implications – With an explicit definition of the roles and responsibilities of the metadata manager, many other digital libraries that need to develop a new or optimize the existing workflow may find the Rutgers experience useful as reference.

Originality – Very few articles have explored this topic although the functions of metadata in the development of digital projects have been talked extensively.

Keywords Workflow, Metadata management, New Jersey Digital Highway, Digital project, Cataloging.

Paper type Conceptual paper

Introduction

Metadata management is an integrated part of project management in the development of digital collections. Because digital libraries have a very short history, metadata manager is a newly emergent title in the library community. The roles and responsibilities of metadata managers have not been adequately defined, thus their collaborative work with digital project team members, colleagues in the library, and/or project partners outside the library, if any, has not been sufficiently described.

Although the collaborative roles of the manager may vary from project to project, the establishment of a common workflow may be necessary for librarians to follow the principles of metadata management, and for people to become familiar with what the manager has done in past projects and how s/he can improve in future assignments. With the rapid development of digital libraries and collections, the importance of metadata management in supporting digitization has become increasingly obvious. In recent years, metadata manager as an independent title has appeared frequently in job descriptions and are in great demand. Some libraries may have reassigned catalogers to work on the new title, while others may ask people with technical backgrounds to manage metadata. It is now necessary to define what jobs a metadata manager should perform so that a standardized workflow can be followed,
which is not only essential to the metadata manager but also helpful to others who interact with the manager at work.

This article will describe the work of a metadata manager in the digital practices of the New Jersey Digital Highway project at Rutgers University Libraries. It attempts to summarize the roles of a metadata manager in a general sense, and at the same time introduce the RUL’s experience in metadata management by presenting some cases. The author hopes that through the outlines of RUL’s experience readers will be able to evaluate the efficiency and effectiveness of an established workflow in the creation and maintenance of metadata and to incorporate the concept into their own practices.

**Literature Review**

In the literature of digitization, the role of metadata in assisting digital project management has been extensively explored (e.g., Hudgins *et al.*, 1999; McCray *et al.*, 1999; Sheth and Klas, 1998). However, very few studies have discussed in detail the role of a metadata manager in the process of carrying out digital projects. In 2002, the UK Intra-Governmental Group on Geographic Information (IGGI) issued the “Principles of good metadata management” to guide the best practices of metadata managers (IGGI, 2002). It calls attention to the importance of managing metadata for information managers and end users and attempts to standardize the management of metadata process. Its limitations are that the compiled principles are for the management of geographic information only and lack sufficient explanations to explicitly define the roles and responsibilities of a metadata manager in handling digital projects, particularly collaborative projects.

It is relatively easy to find studies on the subject of project management in the literature of library and information science (e.g., Burich *et al.*, 2006; Kinkus, 2007; Uwadia *et al.*, 2006). Further studies have made efforts to apply the principles of project management to metadata management and even utilize appropriate tools in the applications (Bennett and Sandore, 2001; Bloss, 2007; Frank, 2007). However, it is difficult to discover articles that focus on the discussion and exploration of the roles and responsibilities of metadata managers, although some articles may have briefly mentioned these as part of the overall description of their digital project undertakings (Bhandar *et al.*, 2007; Bishoff, 2000; Brancolini *et al.*, 2000; Calhoun, 2007; Colley, 2001; Kurth *et al.*, 2004; Lytle, 2006; Murphy, 1998; Wright, 2004). Among others, Kurth *et al* emphasize the necessity of regulating metadata management as a special type of management for data and information and how to redesign the management of MARC metadata in a digital environment. Bhandar *et al* concentrate on describing how social capital can be influenced as a motivator, an integrator, and a facilitator and what its role is in the management of collaborative projects. Both Brancolini *et al* and Bishoff use digital project cases to introduce the work of metadata managers in creating and administering metadata elements. And Lytle has a more detailed coverage on each step of metadata planning, creation, and management in a collaborative digitization project she participated in.
People may think that the definition has been too simple to be made, or the history of managing metadata has been too short and relevant experience too few to produce any generalization. Moreover, very few discussions, if any, have gone any further to present a workflow of metadata management so that the metadata manager’s role in collaborative digital projects can be better understood. In recent years, metadata manager as a position, though may be named differently such as metadata librarian, coordinator, architect, or specialist, has been popular in the library and information community (Coleman, 2005; Simser, 2003). It is now necessary to start defining the roles and responsibilities of the metadata manager.

The New Jersey Digital Highway

The New Jersey Digital Highway (NJDH) is a digital library developed by a grant from the Institute for Museum and Library Services (IMLS) in 2003 [1]. Hosted at Rutgers University Libraries (RUL) – Rutgers, the State University of New Jersey – this digital library is a collaborative statewide project with which a group of New Jersey cultural heritage institutions worked together to build a user-centered information repository aimed at preserving the past and providing digital access for the present and future. As a “one stop shop” for the state’s history and culture geared through photographs, books, documents, periodicals, three-dimensional objects, audio and video, etc., NJDH contains the collections of many libraries, museums, archives, and historical societies. The principal grant partner organizations include Rutgers University Libraries, New Jersey State Library, New Jersey Division of Archives & Records Management, The New Jersey Historical Society, and American Labor Museum/Botto House National Landmark.
Since 2003, RUL has done a tremendous job in making NJDH an attractive repository by providing necessary technical consultations to those that demand. RUL has also made continuous efforts to encourage more participants to engage in the preservation activities, including not only the generous contributions of cultural agencies, but also the active involvement of New Jersey educators in the development of lesson plans toward building upon the digital collections around the state. In doing so, NJDH enables participating institutions to create a dynamic online portal for their collections. This function, giving look and feel of a custom portal so that each institution can organize their own Website interface whereas the digital content is actually linked back to NJDH provides the possibility of centralized database management but at the same time promotes the diversity of digital data presentation.

In order to implement and maintain the operation of the NJDH system, RUL has invested essential technical components to fulfill the system requirements for necessary hardware and software. The investments were taken based on emerging national and international standards to create digital artifacts that stand the test of time, not only for current users but also for future generations. To develop a balanced infrastructure of the digital repository, librarians and staff members with different skills and duties were asked to contribute to the routine work and were organized in such a way that they could collaborate with each other efficiently and effectively inside the library. A feasible workflow was established, as well as improved upon changing working situations, to supervise the undertaking of the projects on a daily basis.

RUL developed a sophisticated software program to provide the flexibility of loading digital objects and inputting metadata elements – the Workflow Management System (WMS). With its online forms connecting to the databases of NJDH, WMS is able to handle metadata creation and maintenance. Since the beginning of its development, WMS has been designed to suit the particular needs of RUL’s digital projects and those of the project partners. In practice, the system has already experienced an on-going refinement over the years so that it can afford more functions to accommodate a wider and more diverse need of digital development and enable online interoperability across domains.
Five types of metadata have been adopted by WMS for the digital collections, namely, descriptive metadata, source metadata, technical metadata, rights metadata, and digital provenance metadata. MODS as the descriptive metadata is convenient to be mapped into the RUL’s online catalog in MARC format, while PREMIS as both the source metadata and rights metadata are better suited for accommodating digital source elements and identifying rights and permissions for individual items. WMS provides controlled vocabulary for LCSH subjects as well as AAT Genre and other standards from which metadata creators can make selections, and also allows the metadata creators to add new terms. WMS has now become a handy tool for the metadata manager in the process of creating and managing metadata standards and elements at RUL as well as with the project partners.

The metadata manager’s role

So far, the most comprehensive description of the metadata manager’s role was made by Weber and Favaro (2007). According to them, this manager “has primary responsibility for metadata design and creation, including working with the project manager to analyze incoming collections, managing metadata timetables and deliverables, periodically reviewing metadata with the collection owner/curator and project manager to insure that the information needs of the primary user group and preservation needs to the owner/curator are addressed, developing templates and controlled vocabularies, training the curator/owner and staff when metadata is created locally.” Their article goes further to describe how the metadata manager has been working with the workflow of WMS in the NJDH projects at the Rutgers University Libraries.

In a nutshell, the assignment of a metadata manager is to plan the selection and integration of metadata standards, coordinate activities involved in the whole process of a project, produce metadata elements, and maintain existing metadata to support data management and information retrieval. Because metadata creation and management are an integrated part of any digital development, the manager will assume extensive interactions with the roles of other colleagues in a project team such as the project manager, the digitization manager, programmers, and assistants. When collaborative projects are taken between the digital team and other organizations, the metadata managers will need to work closely with collection owners of these organizations, either external or internal, to balance resources and keep up an efficient workflow.

Metadata manager as a planner

For any digital project, planning is always the most important part in terms of metadata design. A good plan anticipates every possible scenario that a digital collection could face in future implementation, avoids any repetitive jobs during the process of metadata creation, and ensures a smooth undertaking of the project. In order to make a workable metadata plan, it is
essential for the metadata manager to get completely familiar with (1) the project: its scope, purposes, and all specifics; (2) the potential audience: their needs and possible behaviors in accessing the collection; and (3) the capabilities of the institution and the partners, including technologies such as the availability of necessary hardware and software, financial resources to support the creation and maintenance of metadata, and staffing to work on metadata assignments.

In regard to getting familiar with the specifics of a project, the following questions may represent common concerns of the metadata manager in the planning of metadata organization for most digital projects:

- Type of the collection: e.g., articles, books, manuscripts, music CDs, videos, photographs.
- Size of the collection: e.g., how many items?
- Copyright status of the collection: e.g., copyright protected or in public domain? Who owns it?
- Existence of item descriptions: if yes, in what forms and structure?
- Special information about the collection: e.g., special requirements, controlled subject headings and genre.

Based on an analysis of the characteristics of the project, the metadata manager will have to take the following into consideration: what formats of digital files are most suitable for the collection that can meet the system requirements? What possible naming convention for the collection and for the files so that the management can be handled well and easy for users to search the items? What is the composition of the project team?

If the above mentioned preparations are nothing different from what the roles of other managers – such as the project manager and the digitization manager – need to have in order to start a digital project, the metadata managers will have to go further to evaluate the requirements for and suitability of metadata standards, or metadata elements if the standards have been previously established. Two basic guidelines should be followed in order to make a better metadata plan: the resource description ought to be consistent and as complete as possible (Coleman, 2005).

In general, the following metadata types are typical to most digital projects and may need to be evaluated at the planning stage (see also Ma, 2004; Weber and Favaro, 2007):

- **Description metadata** that provides sufficient information for users to discover, obtains access to information resources, and describes displayed items;
- **Source metadata** that describes the provenance, condition and preservation of analog source materials such as photographs, books, maps, etc.;
- **Technical metadata** that provides information about the digital master files that the system will maintain for long-term preservation and access;
- **Rights metadata** that identifies the rights holders for each information resource and identifies the permissions for use that the rights holder has granted, including any restrictions on those permissions;
- *DigiProv metadata* (provenance) that provides a digital “audit trail” of any changes to the metadata.

At RUL, this evaluation and selection had been made when WMS was developed so that individual digital projects can just follow what WMS has had in its metadata structure. However, it has been understood that each project has its own unique requirements for metadata, and WMS was created to provide the flexibility of mapping metadata elements.

The importance of planning metadata for digital projects has been highlighted in the practices of many other institutions. For example, the Pennsylvania State University Libraries found that using a checklist to identify basic issues for metadata and metadata workflow is key to the success of their digitization projects (Ma, 2004). Ma suggested four steps of preparing and implementing metadata: choosing the schema, deciding record types, creating metadata, and integrating metadata into the online catalog or other database. “Metadata operations would grow more integrated and effective if library metadata managers were to develop metadata processes that promoted sharing and reuse in order to meet the operational needs of the library staff who create and manage digital collections.” (Kurth et al., 2004, 159)

It is obvious that the role of metadata manager as a planner has been widely recognized in the library community.

**Case: The Paterson Chronicle Journal Project**

The *Paterson Chronicle* is a weekly journal for the City of Paterson in New Jersey. Each issue has 16 pages containing many columns such as news, marriage announcements, short stories, TV listings, etc. As many as 328 issues were donated by the publisher with exclusive copyrights to RUL. To bring them online, a project team was organized that included a project manager, a digitization manager, two metadata managers, two serial cataloging librarians and two metadata assistants. Because this was the first project of a series of journals in NJDH, the initial discussion of the team concentrated on the structure of the materials being organized. The team believed that a well-planned structure could make data retrieval and management an easy job. For example, the series could be organized under the hierarchy of collection → volume → issue, or arranged simply by sequential issues. In terms of visual presentation, the series could be organized by individual images to represent each page of any issue sequentially, or be systematized with a structure map that served as a table of content for each issue, from which navigation between individual images (pages) would be simple. Full text searching was designed with the assistance of the optical character recognition (OCR) software.

What the metadata managers needed to judge, in addition to the above considerations, was the types of information to be selected for metadata. Usually, like journals will have “filler” or “junk” to fill in empty spaces that may be less relevant to the journals and may need to be neglected in metadata creation. Nonetheless, local information in the journals will be the most relevant information and is important for metadata selection. This understanding was consulted with and confirmed by librarians and historians.
The preparation of the controlled vocabulary was equally important in the planning of metadata elements. Typical concerns in the preparation phase included an analysis of the journal title and the selection of personal names in the records. During the past 12 years, the journal title has been changed four times, each of which contained the word “Chronicle”. Hence, the metadata managers kept asking the following questions: Is it better to use only one title instead of four different titles so that users will not get confused? Do we need to add the other three names as other titles? Do we also need to add a qualifier to the title such as “Paterson, N.J.”, because “Chronicle” is a popular journal name, if we choose this word as the journal title?

In regard to personal names in the metadata creation, the metadata managers felt the need of balancing the time and energy of team members and the possible requirements of future data search for the journals. Obviously, there have been too many personal names available in the journals, and the inclusion of all of them was both impracticable and unnecessary. For example, a movie advertisement may contain the names of film stars who were not related to the locality. Yet, the question was: what are the criteria to select names for the metadata? To answer the question, it was important to imagine the potential user group – historians who are interested in local history and genealogists who may want to look for information of their ancestors or friends. Therefore, the metadata managers determined that local politicians, community leaders, and companies would have their names searchable, so were names of people who were the subjects of the journal articles and photos. For example, both the bride and groom of a marriage statement would appear in the metadata, but not their parents.

**Metadata manager as a coordinator**

Virtually every digital project is a collaborative project. In most institutions, digital projects are carried out in the technical service division, while the collections being digitized and processed belong to another part of a library, mostly special collections. Metadata managers do not usually know the collections. Therefore, it is the responsibility of the metadata manager to work closely with curators or librarians of these collections who are acquainted with their materials and users.

When a collection belongs to an organization outside of the library, or even outside the institution with which the library is affiliated, collaborations become even more important. Such external collaborations are also more difficult to take. NJDH is a collaborative venture between RUL and other New Jersey cultural and educational institutions with partners from different parts of the state. Many of the projects in NJDH are external (Rutgers) collaborative projects and required a lot of communications between the metadata manager and people in the host collections. Some collaborative projects are even from outside the state. For example, RUL and the Library of Virginia Polytechnic are collaborating on a project that aims to build a digital repository for the Virginia Tech memorial collection dedicated to the tragedy that occurred on campus in April 2007.
This coordination role can be also extended into the project team within which relationships between different roles of team members ought to be maintained strongly. A typical digital team at the NJDH may include a general Project Manager, a Collection Manager/Curator, a Metadata Manager, a Digitization Manager, an Objects Creator, and a WMS Manager (Weber and Favaro, 2007), although the composition may vary.

Case: The Hoboken Historical Photographic Project
This is a collaborative project between RUL and the Hoboken Public Library (HPL) for a collection of more than 1,000 photographs taken during the 1800s to the 1900s. HPL is a small library without necessary equipment, expertise, and staffing for digitization. In order to carry out the collaboration, the HPL Director assigned two librarians to work on the project. At RUL, a team was organized to include a project manager, a digitization manager, and a metadata manager. The team worked closely with these two HPL librarians as well as with RUL supervisors – the Associate University Librarian for Technical Services and the Director of the Cataloging and Metadata Department – and RUL cataloging and system librarians.

It was the responsibility of the metadata manager to coordinate the planning and creation of metadata for the project. In the planning, the metadata manager, upon discussions with the supervisors and team members, decided to apply a batch-loading procedure to edit and transfer the existing textual descriptions of the collection at HPL into WMS. This would significantly save the time and energy of metadata construction. Yet, the automation of metadata processing was new to everyone, including the project team members who were the key persons to the success of the digitization. Therefore, the first task of the metadata manager was to train the team members the automation process.

During the implementation of the batch loading, it was important that the metadata manager kept close communications with the librarians at HPL. The loading process was taken by following a workflow of “evaluation of the collection descriptions → communication with the host librarians → editing the descriptions into metadata formats → testing the metadata → implementation of the batch loading into WMS.” The steps can be repeated according to actual work and requirements, and the communications between the metadata manager and the HPL librarians was crucial to ensure the appropriateness and correctness of the metadata elements. Both email/telephone contacts and physical visits were made throughout the project cycle to exchange ideas and findings and to solve problems encountered.

Metadata manager as a facilitator

If the role of the metadata manager as a planner is building a framework for the development of metadata, and the role of the manager as a coordinator is organizing activities and communicating with other roles, the role of the manager as a facilitator is more about actually creating metadata elements. The facilitator role may vary among organizations or even
between individual projects. Yet, in most cases, the manager may actually participate in the inputting of metadata elements, or supervise others in the inputting process. The supervision role may be taken to direct assistants to create metadata elements. Many libraries hire student assistants or ask cataloging staff to do the actual metadata creating job. Because the creation of metadata is always a tedious job that requires special care and good understanding of the project being taken and the metadata being made, a close supervision is necessary in addition to initial as well as continuous training. It is relatively common that the metadata manager constantly discovers mistakes in metadata fields input by assistants due to their fatigue or inadequate knowledge of their work.

When collaborative projects are taken, staff of the partner organizations may participate in metadata creation. With little or no knowledge of metadata, they also need training and supervision. An experienced metadata manager knows such supervision has been more difficult for partner metadata creators than for internal assistants who are in close proximity and easy to communicate with. With different types of digital projects in the NJDH endeavor, most of them involve the contributions of metadata creation from the partner side.

In processing metadata, the manager may be handling hand-input of metadata elements, or managing automated batch loading of existing data into metadata fields. The former task is usually taken with the assistance of online templates created exclusively for metadata creation. The latter task requires additional jobs that may include the whole process of evaluating and editing existing data, mostly textual descriptions of the collection items, testing the results of initial batch-loading, re-evaluation, and final loading. The evaluation and editing of the original data is key to the success of a batch load, which may require reorganization of the original data with the possibility of combining and splitting data fields. The metadata manager facilitates batch loading.

**Cases**

Four people (two metadata managers and two assistants) worked together to hand-create metadata for the Paterson Chronicle Journal project. It was required that the metadata must be consistent throughout time. At the design stage, it was agreed that a structure map was good to serve as a table of contents for image-level navigation. In the actual construction, a variety of trivial matters surfaced and waited for solutions. For example, in terms of metadata creation, it would be problematic to use an image to simply represent a journal page, even if this image was an exact digitized version of the page. This page may not be indexed in the original journal’s table of contents, or was paginated differently as the image. Careful work on details became an everyday routine of the metadata team.

With regard to batch-loading metadata elements, tedious work by hands would be replaced by the assistance of software programs, which was the strategy RUL adopted for the metadata creation of the Hoboken Historical Photographic project. However, choosing batch loading was by no means to disregard carefulness on metadata creation. Instead, the metadata manager needed to spend a lot of work on evaluating and editing the textual descriptions of the collection so that the data could be accommodated properly into the metadata fields. The
original descriptions were formatted in a spreadsheet by similar fields, but not necessarily matchable, to RUL’s metadata fields, requiring extensive efforts to make them exactly the same as the fields in WMS. Several operations were taken to fulfill the task of the mapping, including merging, splitting, combining, and simply transferring of original datasets into new metadata. Any mistakes in the work would make future data transferring unsuccessful. For example, the “&” sign could not be correctly handled by WMS and caused problems for batch loading. Improper reorganization of the datasets, if successful in the transferring, would make data retrieval ineffective or, even worse, impossible.

**Metadata manager as an administrator**

As an administrator, the metadata manager is responsible for setting up different levels of accessibility for the database management system for people who need to work on metadata. According to their assigned privileges, participants may have at the super user level an unlimited access to all resources, at the organization level an access to all collections and projects owned by the assigned organization(s), at the collection level an access to all projects belonging to the assigned collection(s), or at the project level an access to assigned project(s) only.

As soon as metadata are created and reserved on the system, they will serve for organizing and explaining digital objects and making online search possible. But, the completion of metadata creation is by no means the end of the work of metadata management. Instead, it requires continuous maintenance of the metadata to reflect the changes of (1) digital files or data structure, (2) system infrastructure, (3) users’ new requirements for a project, (4) metadata standards, and the like. The role of metadata manager as an administrator is to take metadata stewardship by checking metadata entries for completeness and controlling metadata integrity so as to ensure that the metadata can support the normal operation of digital collections in data organization, preservation, and retrieval. The manager needs to align with key stakeholders (e.g., power users and collection owners) and to advise on process optimizations in the interaction between these roles.

**Conclusion**

The roles of a metadata manager as a planner, coordinator, facilitator, and administrator on collaborative digital projects have been outlined in this article. His or her responsibilities and practices in structuring, capturing, management, dissemination, manipulation, and preservation of metadata focusing on new and existing data in support of digital collections have been briefly described. It is worth noting that the different roles and responsibilities of the metadata manager are not isolated with each other in the undertaking of any digital project. Rather, the manager needs to iterate the roles and responsibilities and work closely
with many other roles in a partnership environment. It is also worth noting that the development and management of metadata is project-independent.

This article is only an attempt to portray a general workflow with the reference of the NJDH experience, particularly on collaborative digital projects. It is hoped that this article can serve by drawing the attention of researchers from metadata management to the discussion and definition of the roles and responsibilities of metadata managers. With theoretical potential and practical significance, such discussion will help enhance peoples’ understanding of metadata development and administration and recognize the importance of transformation processes to library operations.

Libraries are not alone in facing the challenges of digital development. Other types of cultural institutions such as museums, archives, and educational agencies, have also participated in the efforts of digitization and online information disseminations. Collaborations among them have become a popular practice. Metadata managers have already played a central role in such collaborative projects. A better understanding of the managers’ role will certainly help facilitate these collaborations. The experience at Rutgers University Libraries presents only one case of such cross-institutional collaborations in digital endeavors, however. More discussions, particularly those from metadata managers as practitioners, on the subject are advocated.

Acknowledgements

The author wants to thank Grace Agnew, the Associate University Librarian at Rutgers, for providing me opportunities to work on NJDH projects and metadata management. I am grateful to Mary Beth Weber and Rhonda Marker who provided great helps in carrying out metadata creation and maintenance.

Note


References


IGGI (Intra-Governmental Group on Geographic Information (2002), The principles of good metadata management.


