Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:05 AM

Current Topics in Library and Information Practice

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:05 AM

Organization, Representation and Description through the Digital Age

Information in Libraries, Archives and Museums

Edited by
Christine M. Angel and Caroline Fuchs

DE GRUYTER
SAUR

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:05 AM

ISBN 978-3-11-033715-0
e-ISBN (PDF) 978-3-11-033741-9
e-ISBN (EPUB) 978-3-11-039599-0

Library of Congress Cataloging-in-Publication Data
A CIP catalog record for this book has been applied for at the Library of Congress.

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.dnb.de.

© 2018 Walter de Gruyter GmbH, Berlin/Boston
Typesetting: bsix information exchange GmbH, Braunschweig
Printing and binding: CPI books GmbH, Leck
♾ Printed on acid-free paper
Printed in Germany

www.degruyter.com

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:05 AM

Preface

As the lifecycle of the web has evolved from the read-only web of 1.0 to the Semantic Web of today, libraries, archives, and museums (LAM) continue to manage and describe their objects utilizing varying degrees of interoperability. However, as the web has continued to develop so to have the organizational and descriptive practices among LAM professionals as they transition from cataloging objects within their own community-specific domains to the practice of linking structured data on the Semantic Web. What follows is a series of case studies written by current practitioners and researchers documenting their experiences in how they are making this transition. Through these shared experiences, information professionals across the LAM environments have the opportunity to learn about the construction of shareable metadata, the implementation of a technological infrastructure and the tools used to facilitate easy discovery and use of this metadata, and organizational restructuring within the LAM environments that support the management of these new technologies that enable one to participate within the Semantic Web.

https://doi.org/10.1515/9783110337419-202

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

Contents

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:05 AM

Part III: Finding Structure. Making Connections. Web 2.0

Part IV: Transition to Web 3.0

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:05 AM

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:05 AM

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:05 AM

Introduction

There is awareness among information professionals today that related information objects can be linked together via their shared data elements within the online environment. However, the actual practice of linking shared data elements among related information objects across the library, archive and museum (LAM) environments have been slow. Part of the reason for this is because data content standards used for the organization and description of information within our LAM environments have only recently evolved to reflect the entity-relationship (Chen 1976) structure of online catalogs used today. While continued revising of data content standards have enabled more granular descriptions for the retrieval of information objects within the online environment it is still not possible for information users to search for related objects across collections. Evidence of this can be found within the various information systems LAM professionals use to manage and provide access to collections within their own organizations.

The main reason as to why cataloging standards practiced within today's LAM environments are not interoperable for the retrieval of objects within the online environment is because the World Wide Web had to evolve through different stages of development in order to accommodate such access (Choudhury 2014; Solanki and Dongaonkar 2016). These stages of development are defined as web 1.0, web 2.0 and web 3.0. Web 1.0 is a "read-only" web that can only be understood by humans but was innovative because it removed the geographic boundaries typically associated with our brick-and-mortar institutions (Berners-Lee 1998; Choudhury 2014; Solanki and Dongaonkar 2016). Web 2.0 was coined by Dale Dougherty and Tim O'Reilly (2005) and its defining characteristic is that users now have the ability to not only consume information but also be producers of it (Friedman 2005). The defining feature of web 3.0, also called the "Semantic Web," is the addition of machine readable documents (Choudhury 2014; Solanki and Dongaonkar 2016). According to the World Wide Web Consortium (2015), "Semantic Web technologies enable people to create data stores on the web, build vocabularies, and write rules for handling data. Linked data are empowered by technologies such as RDF, SPARQL, OWL, and SKOS" (first par.).

It is with the development of the Semantic Web that information professionals now have the technological tools available to provide information users access to related objects across collections within the online environment. However, the majority of practitioners charged with managing and providing access to cultural heritage information have yet to integrate the technological tools or

https://doi.org/10.1515/9783110337419-001

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

"...set of best practices for publishing and connecting structured data on the web known as Linked Data" (Bizer, Heath, and Berners-Lee 2009, 1). "Technically, Linked Data refers to data published on the web in such a way that it is machine-readable, its meaning is explicitly defined, it is linked to other external data sets, and can in turn be linked from external data sets" (Bizer, Heath, and Berners-Lee 2009, 2). While the integration of Linked Data (LD) principles have not been fully integrated into the practitioner's organizational and descriptive practices, "the library, archive, and museum (LAM) communities have put considerable work into developing new LD tools, standards and published vocabularies, as well as explored new use cases and applications" (Mitchell 2016, 5). Evidence of this exploratory work can be gleaned from the case studies presented in this book as researchers and practitioners begin to construct shareable metadata, implement new technologies, and restructure their organizational environments to support the management of these new technologies.

What follows is a selection of case studies from leading practitioners and researchers, each representing significant concepts or themes concerning the organization and representation of information. All case studies presented in this book lend value as researchers and practitioners describe their experiences in adopting new technological tools, organizational frameworks, and descriptive practices as the lifecycle of the web has evolved through its different stages of development. The book begins with part one which provides a brief overview of the historical development of pre-web cataloging technologies and discusses their influence on the organization and description of information as it is found within libraries, archives, and museums.

With the adoption of web 1.0 technologies organizational and descriptive practices within the library, archive, and museum environments changed (Svenonius 2000). Experiences documenting how different practitioners responded to these changes in technology are offered in part two. The main themes presented are concerned with providing access to cultural heritage materials while maintaining the integrity of collections within the online environment. Part three provides a snapshot of different experiences current practitioners have gained in organizing, describing, and making information accessible utilizing participatory web 2.0 technologies. Part four of this book consists of case studies representing information organizations that are currently finding ways to aggregate collections from a variety of institutions with an eye towards participating within the Semantic Web.

Chapter one begins with Michael Handis providing a historical analysis of the development of libraries, archives, and museums and the access tools used within each to organize and retrieve information. The evolutionary concept of the catalog from medieval and modern times is introduced as is the purposes of

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

them. Handis describes several elaborate organizational systems created over many centuries in different parts of the world. He also documents a variety of different media types that were used for the recording of information throughout the ages. The catalog is defined as a common tool that aggregates information about a collection, which is key to discovering the holdings of that collection. Handis also differentiates among the different types of catalogs where finding aids serve as an in-depth listing of what is in a specific collection and museum catalogs, known as registrars, which contain a plethora of information such as accession number, donor, and object history. Handis concludes by stating that while the organization of information can take many forms it must make sense to the user.

Part two of this book offers a series of case studies different practitioners and researches have experienced in moving their collections from the pre-web to the web 1.0 environment. In chapter two, Michelle Levy and Christina Orozco describe the various stages they went through in migrating archival documents from analog to digital format. Because a number of staff working in the offices of The Missionary Society of St. Paul the Apostle were still accustomed to paper documents, the author's attempts of applying web 2.0 technologies to the digitization of the archival documents did not work. As such, their solution was to adapt to the organization's specific web 1.0 administrative needs while leaving pathways open for the collection to ease into web 2.0 in the future.

The case study presented in chapter three offers a case study describing a project practitioners from the Texas Tech University School of Music (SOM) embarked upon in order to transfer their classical music collection from their SOM Listening Library to an online system. According to the authors of this chapter, the SOM Listening Library offered content on physical media which was circulated to its patrons via Compact Disc (CD). Transferring the classical music collection to an online system would increase access to the collection. However, the streaming-sound collection project failed because (1) it was too difficult to use, and (2) the music streaming industry had made the streaming-sound system obsolete. Many lessons were learned during the course of this project such as avoiding the "sunk cost fallacy" and to build in assessment strategies. The beauty of this case study is grounded in the fact that the authors hope to help others to better plan and execute innovative projects and metadata practices at their institutions.

Dara Lohnes-Davies, Manager of Collections at the University of Wyoming Art Museum, describes best practices implemented during a digital enhancement initiative of cultural heritage objects. These best practices were developed during a two-year grant-funded digitization project while working as the Photographer at the National Music Museum (NMM) located in Vermillion, South

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

Dakota. The goal of the digital enhancement initiative was to digitize 3,357 musical instruments belonging to Arne B. Larson thereby making the collection accessible within the online environment. This involved processing photographs, creating two-dimensional (2D) image packages for each musical instrument, and making the information available for online consumption along with a website overhaul. As such, this chapter focuses on digitization efforts and photography workflow processes the NMM staff implemented in completing their two-year grant. By presenting this information as a case study, the staff of National Music Museum (NMM) hope to help similarly sized institutions plan their own digital access enhancement initiatives.

Tess Colwell, Digital Asset Associate at the Brooklyn Historical Society (BHS), describes how staff members have refined their management and decision-making processes in the digitization and description of objects held within their collections. One of the main difficulties in making collections available to the public at the BHS has to do with the management and processing of collections. Prior to 2009, collections were separated by object type. For example, if a collection included both archival manuscripts and photographs, staff members would remove the photographs from the collection and process them separately from their related archival documents thereby destroying the integrity of the collection. To rectify this, the BHS has redefined its processes for digitization and description and are now trying to reestablish the integrity of their collections. Utilizing the concepts of "More Product Less Process" and the "Accessioning as Process Model," chapter five describes how the BHS has refined their management and decision-making processes in the digitization and description of objects held within their collections.

As web 2.0 technologies increased in use within the library, archive, and museum environments the adoption of new cataloging tools was essential in order to accommodate the new interconnections between people and information (Marty and Jones 2012). For the first time information users are becoming participants in describing cultural objects, a task which was traditionally held by the practitioner. The information user can now participate in the construction of new knowledge and contribute to the interpretation of museum objects and exhibits. New interconnections between information and technologies (Marty and Jones 2012) were also experimented with as information professionals participated in shared cataloging standards and attempted to maintain the hierarchical structure of their collections.

According to Jane Zhang, Assistant Professor within the Department of Library and Information Science at Catholic University of America states that within the archival environment, archivists focus on representing the context and content of an archival collection as an organic whole thereby integrating

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

them into one hierarchical structure. However, with the movement towards digitization, archival documents are converted into digital objects and assigned item-level metadata. Zhang questions how archival context will be preserved within the digital environment by assessing three digital archival collections, each consisting of a unique digital representation approach. The three different approaches to the organization and description of archival materials within the online environment reveled various levels of archival representation in terms of context, digital content accessibility, and interoperability.

In his chapter titled *Linking Items, Connecting Content: The Donald Thomson Collection*, Mike Jones argues that aggregations are based on networks of connections perceived between things rather than the inherent qualities of the things themselves. As such, the relational data connecting one object to another within a collection allows for the complete interpretation of the total collection. This is what makes objects within a collection and the relationships between them discoverable. However, within the digital world, such documentation is limited due to the rigid design of online databases. Jones concludes stating that in order to document relationships between and among objects within an online collection practitioners need to consider the representation of the collection of networks and the webs of meaning between objects thereby providing a richer information environment to users and more effectively sharing the knowledge held by museums with the broader public.

In chapter eight, Jennie Choi and Giovanna Fiorino-Iannace discuss the evolution of the use of technology within the museum environment with the introduction of an image database at The Metropolitan Museum of Art (The Met). The introduction of an image database at The Met began over twenty years ago with the unification of the museum's textile collections into one resource center. Known today at the Antonio Ratti Textile Center, this collection represents a turning point in The Met's history because it changed the way catalog information was accessed by the museum staff as well as the way it was presented to and accessed by the public audience. Couched within the goal of providing accurate descriptive information about The Met's extensive collection of art as an ongoing process which has evolved from using pre-web technologies such as typed catalog cards at the beginning of the twentieth century to the development of computer-based platforms representative of the web 2.0 environment.

Today, The Met maintains one central collections management system that combines all data relating to the basic description of every object in its collection, as well as representative images of each object, and information on historical relevance, current condition, and storage location. This chapter begins with a discussion of the evolution of the use of technology at The Met with the first part exploring how the creation of the Antonio Ratti Textile Center, The Met's

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

main repository for the storage and study of its historical textile collection, paved the way for the eventual implementation of one consolidated database for the entire art collection. Choi and Fiorino-Iannace conclude with how the transition to one collections management system for the entire museum required metadata standardization and data formatting along with the creation of one institution-wide cataloging standard.

Layna White, Head of Collections Information and Access at the San Francisco Museum of Modern Art (SFMOMA), describes the organizational and descriptive practices of art documentation using a surveillance-based installation that has been exhibited three different times titled *Predictive Engineering*. Charged with maximizing access and sharing of the museum's digital records for staff and public benefit, White utilizes the surveillance-based installation as an opportunity to explore information documentation within the museum environment, with the museum visitor acting as coauthors in the description of works of art. Such documentation is essential, White argues, because the questions posed by museum visitors are central in describing a work of art. White explores this idea through the use of wikis that can be used as a space for museum visitors to act as coauthors in the documentation of art exhibits. In this way, wikis are used just as archival documents and library collections when describing museum exhibits and works of art. The chapter concludes with a discussion on the limitations data content standards and the tools used to management them serve the needs of art documentation within the museum environment.

In chapter ten, Emily Griffin and Rachel Lipkin discuss their experiences in using an open source geographic information system (GIS) tool used by the New York Public Library (NYPL) called the MapWarper portal. The portal was created in order to provide user's access to the NYPL's Firyal Map Collection, which consists of over 10,000 maps depicting New York State and surrounding areas from the sixteenth to the nineteenth centuries. The maps depict historical information about the state's transportation lines, real estate development and topographical information. With the MapWarper tool, patrons are able align points on historical maps with corresponding points on real-time maps, allowing data from the present day to be layered over historic maps. As such, this chapter examines issues surrounding how GIS tools are affected by different metadata schemas and how the data protocol utilized by the NYPL affects the usability of the MapWarper portal.

Within the machine readable environment of the Semantic Web, the principles for the creation of linked data (Tim Berners-Lee 2006) are slowly beginning to emerge within the organizational and descriptive practices of our library, archive, and museum environments. Evidence of this can be gleaned from the

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

case studies presented in part four of this book as practitioners and researchers share their experiences in the construction of shareable metadata, the implementation of a technological infrastructure that enables easy discovery and use of this metadata, and organizational restructuring within the LAM environments that support the management of these new technologies.

The construction of shareable metadata is a necessary component of the Semantic Web because "Linked Data should be published alongside several types of metadata, in order to increase its utility for data consumers" (Bizer, Heath, and Berners-Lee, 2009, 7). Additionally, "systems that consume linked data must evaluate quality and trustworthiness of the data. A common approach for data quality assessment is the analysis of provenance information" (Hartig 2009, first par.). In other words, for an information user to be able to assess the trustworthiness of a resource, "data should be accompanied with metadata information about its, creator, its creation data as well as the creation method. Basic provenance meta-information can be provided using Dublin Core terms or the Semantic Web" (Bizer, Heath, and Berners-Lee 2009, 7).

Practitioners are beginning to question the trustworthiness of their metadata creation methods due to the high degree of interpretive work during the cataloging process. In his chapter titled *Transcribe as Seen – Challenging RDA Regarding Gender in Moving Image Materials*, Wagner argues that Resource Description and Access (RDA) rule 9.7 is not flexible enough to express current understandings of queer and transgender identities. His premise is that by assuming a categorical binary of gender completely disregards how a person identifies on a spectrum outside of this binary. According to Wagner, this limitation inhibits the information user's ability to retrieve resources concerning queer and non-gender conforming materials. In chapter eleven, Wagner challenges the functionality of gender within RDA cataloging practices and provides suggestions for alternative ways in which RDA rule 9.7 can be expanded for more inclusivity.

Sonia Yaco, Assistant Professor and Special Collections Librarian at the University of Illinois at Chicago present a methodology for the construction of shareable metadata. In their chapter titled *Cultural Heritage Curriculum Crosswalk: Using Metadata to Connect Curriculum*, Yaco, Rizvi, and Ramaprasad present a new framework to link curriculum to collections through the creation of an ontological metadata crosswalk. The benefit of creating such an ontology is that it helps cultural heritage professionals, university administrators and systems decision-makers to speak a common language and reduce ambiguity to set desired outcomes. The end result is that the ontology will help in the development of a software solution to automate matching, which will allow collection metadata exploitation to enhance the value of collections for a curriculum

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

and will allow curriculum to drive the utilization of the collections. Chapter twelve begins with a review of literature, then the methodology used in developing the ontology is presented and concludes with steps for turning the ontology into practical solutions which can be applied to any type of cultural heritage collection.

With an eye towards adopting principles of linked data (Berners-Lee 2006; Bizer, Heath, and Berners-Lee 2009), practitioners from the University of Nevada, Las Vegas (UNLV) document their experiences in reconceptualizing their metadata management practices. The lesson learned from this experience was the need to rethink project workflows in order to manage metadata practices for the creation of an interoperable environment. Other practitioners, such as those working at the University of Virginia (U.Va.) Library, recognize that while there is a need to rethink metadata project workflows this task has been difficult to implement as the conceptual role of metadata librarian has been ambiguously defined since its introduction as a position in the late 1990s.

In chapter fourteen, Ivey Glendon, Manager for the Metadata Analysis and Design department at the University of Virginia (U.Va.) Library, discusses how the institution has arrived at its contemporary understanding of metadata librarianship within today's organizational structure. In this chapter, Glendon analyzes metadata librarianship within the context of major organizational changes that have taken place in the past decade at the University of Virginia Library. The chapter concludes with examples of how the Library's newest conception of organizational design around metadata work has enabled significant advancement in the Library's ability to meet new and changing university directives.

According to Jocelyn Olson Triplett, Manager of Digital Content Management and Dissemination at the University of Virginia (U.Va.) Library holds the premise that in order for one to increase the interoperability of their collections it is necessary to first determine the level of interoperability needed in order to link to collections across institutions. As such, in chapter fifteen Triplett discusses the challenges encountered while standardizing and migrating several varieties of metadata, the technological architecture implemented by the U.Va. Library for dissemination of the collections, and how decisions made during the migration process affected the discovery and display of the collections in U. Va.'s online public access catalog. The reward, according to Triplett, for completing collection migrations is the consolidation of all metadata for image collections into a standard metadata format repository that increased the quality and retrieval effectiveness for the discovery of relevant records within the standardized collections. Additionally, by selecting a nationally supported metadata schema and consolidating the collections' metadata into one reposi-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

tory collection, the U.Va. Library was able to aggregate their records for the Digital Public Library of America (DPLA) thereby making their information resources interoperable with collections across other institutions.

In an effort to improve the quality of metadata for its digital collections and operate within an aggregated digital environment, the University of Mississippi Libraries (UML) embarked upon a massive metadata cleanup project of its existing collections in CONTENTdm. While the adoption of nationally recognized metadata standards and best practices were always considered important to the UML, the decision to join the Digital Public Library of America (DPLA) provided the impetus needed for this massive undertaking. There were a number of tasks that took place during this cleanup process such as the creation of metadata guidelines, implementing OpenRefine to clean the metadata prior to ingest, and to develop a workflow plan for the creation of new metadata. Of particular note in this case study was the decision to consider the future implementation of linked data. Two key lessons learned during this massive cleanup project was that metadata requires continuous maintenance and that if the UML is going to continue to participate in large scale projects such as the DPLA national library then they need to think about their metadata within a broader context.

In chapter seventeen, Samantha Norling, Archivist at the Indianapolis Museum of Art (IMA) Archives describes the processes and tools used when establishing connections between the Museum Archive and the Museum's permanent collection. In order to merge the different metadata standards, data elements employed by the museum registrars and curators were mapped to related data elements used by the archivists. Results of this aggregated metadata mapping process culminated into the creation of a metadata mapping tool. This tool made it possible for the IMA Archives to be the very first institution that was not using CONTENTdm to contribute to Indiana Memory, which is a statewide Service Hub for the Digital Public Library of America (DPLA). The tool now serves as a test case for the new DPLA Indiana Metadata Mapping Tool.

Charged with a mandate to enhance learning, the Drexel Digital Museum (DDM) project provides a platform for the information user to travel through simulated galleries and manipulate contextualized three-dimensional (3D) objects. In this final chapter, the authors describe the DDM project as an international multidisciplinary collaboration of researchers focused on the creation of new media for the exhibition of cultural heritage objects. Through these collaborative efforts, research on classification structures and metadata mapping revealed the development of new standards and strategies one can employ for sharing information among different databases. Additionally, by using an object oriented approach, the authors describe the tools and processes used for the creation of permanent links between the 3D information object and its re-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

lated metadata. The DDM's next goal is to eventually release this collection as linked data within the Semantic Web.

References

Berners-Lee, Tim. 1998. "The World Wide Web: A Very Short Personal History." https://www.w3.org/People/Berners-Lee/ShortHistory.html.

Berners-Lee, Tim, 2006. "Linked Data-Design Issues." https://www.w3.org/DesignIssues/LinkedData.html.

Bizer, Christian, Tom Heath, and Tim Berners-Lee. 2009. "Linked Data – the Story so Far." *International Journal on Semantic Web and Information Systems* 5 (3): 1–22. doi:10.4018/jswis.2009081901.

Chen, Peter Pin-Shan. 1976. "The Entity-Relationship Model – Toward a Unified View of Data." *ACM Transactions on Database Systems* 1 (1): 9–36. doi:10.1145/320434.320440.

Choudhury, Nupur. 2014. "World Wide Web and Its Journey from Web 1.0 to Web 4.0." *International Journal of Computer Science and Information Technologies* 5 (6):8096–8100.

Friedman, Thomas L. 2005. *The World Is Flat: A Brief History of the Twenty-First Century*. New York: Farrar, Straus and Giroux.

O'Reilly, Tim. 2005. "What is the Meaning of Web 2.0?" http://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html#mememap.

Solanki, Manishkumar and Abhijit Dongaonkar. 2016. "A Journey of Human Comfort: Web 1.0 to Web 4.0." *International Journal of Research and Scientific Innovation (IJRSI)* 3 (9): 75–78.

World Wide Web Consortium (W3C). 2015. "Semantic Web." https://www.w3.org/standards/semanticweb/.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

Michael W. Handis

1 The Historical Use of Catalogs in the Arrangement of Knowledge in Libraries, Archives, and Museums: A Survey

Introduction: The library, archive, and museum environments

"It is impossible, in fact, to attach too much importance to the advantages resulting from an intelligent and methodical order in the arrangement of a library. Of what utility would be the richest treasures if it were not possible to make use of them" (Cotton des Houssayes 1906, 44)? This is also true of archives and museums as well as libraries. It is only after cultures create written language do they then need to create repositories to collect and hold the information created. Libraries and archives came into existence around the same time in the differing histories of the world. Museums, too, are the products of advancing cultures, which seek to celebrate civilization's cultural accomplishments. The catalog, the assembling of information about a collection and the key to discovering the holdings of that collection, is a common tool: "Throughout history, library catalogs served a variety of purposes. From inventory lists of valuable property to marketing tools designed to solicit donations, the library catalog has played a varied role in serving both the library and its users" (Clarke 2014, 1).

Not only libraries, but also archives and museums use catalogs to document collections and donations. Archives have finding aids, but usually the information about collections are in a catalog. The finding aid is usually an in-depth listing of what is in a specific collection, with the summary information being included in a catalog. Museum catalogs contain a plethora of information, if exhibition catalogs are any indication: the accession number, donor, history of the object, area from which it came, what it is, etc.

The term "catalog" will be used in the broadest sense in this chapter, meaning the recording of individual items in a collection (for libraries, a book list), breaking those items into broad fields of knowledge, and then as a way of giving access to a collection through author, title, and subjects. The assembling of even basic information is the first step in the creation of the catalogs that still survive from medieval and modern times.

https://doi.org/10.1515/9783110337419-002

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

There are a variety of ways that collections of materials can be organized. Consulting the finding aids for archival collections is often the only way to confirm what materials are held in the collections. Public libraries organize their collections differently from academic and special libraries. Public library patrons are expected to browse the various individual collections, such as travel or local history, shelved separately from the main collection. In contrast, academic and special library patrons primarily use the collections for research, using the library catalog to search for what they need. The Dewey Decimal Classification System (DDC) and the Library of Congress Classification System (LCC) are different catalog numbering systems, but both create shelf organization by subjects and authors. Such in-depth organization is a result of mass literacy and print publications, producing more and more materials that needed an ordered organization to make them easier to find. Such detailed systems of organization are only needed when a collection is large, containing perhaps 1000-plus titles. Nonetheless, there were several elaborate systems created in different parts of the world over many centuries.

Sometimes the organization of a collection seems to defy logic, such as the Jagiellonian Library in Cracow (founded 1364), which organized its collections to honor the benefactors of the institution and not by subject or author (Nowak 1997). Of course, the central question is, what was the purpose of this catalog? Was it used to promote the library, or to honor the donors so that they might make future donations? This practice of adhering to the way a donor arranged his/her collection was not necessarily the best organization for other patrons who might use the library. In the case of the Jagiellonian Library, keeping donated materials in their original organizational schema meant that multiple catalogs needed to be consulted for use of the library collections, since each collection would have its own catalog. It is possible that an integrated catalog, one catalog listing all holdings in the library regardless of what collection it was in, could have been created but with the technology of the time it was not easy to create such a catalog let alone maintain or keep it up-to-date.

The first library catalog created in China was for the Han Emperor Cheng (reigned 33–7 BCE), who wanted to know what materials were held in the imperial library. It took scholars nearly two decades to organize the collection and record it in a catalog that was presented by the son of the original compiler to Cheng's successor. Though this first catalog was created simply to tell the emperor and the court what was in the imperial library and not an attempt to document and organize all of Chinese literature up to that time, this cataloging system influenced the way materials were cataloged in China up until the beginnings of the twentieth century, when Western cataloging practices were introduced.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

Athens, the Hellenistic successors, the *Pinakes*, and Rome

Libraries and archives evolved around the same time in world history, and which came first is impossible to determine. Museums existed in the ancient world, but it appears they came after libraries and archives. In ancient times, a library was more than a repository for books. Libraries served as scholar-centers and, later in the Roman Empire, as public areas for events. Archives housed primarily those documents that were generated by the governments of the time to be used as historical records and materials to consult when governing. Museums at this time were primarily outdoor, "... the public monumental portico-temples of Rome as museums where public collections of art were exploited by ancient generals, politicians and emperors for political ends" (Macaulay-Lewis 2009, 2). The Greeks also had parts of their gymnasia, the *exedra*, where people sat and watched the athletic competitions while amongst displayed works of art. As will be discussed, these great outside areas for art were started during the Hellenistic Age (323–30 BCE).

According to Dorothy May Norris, the Babylonian library built at Akkad/Agade by Sargon I in the seventeenth century CE had a classified catalog organized by subject, which gave instructions on how to find the books (Norris 1969). However, recent scholarship by Eleanor Robson into ancient Babylonian and Assyrian libraries has found little or no evidence that these libraries had librarians, catalogs or any type of finding aid (Robson 2013). She attributes this to the radically different functions that Babylonian and Assyrian libraries played as opposed to Roman libraries, specifically the practice of restricted users, with no dedicated library buildings (libraries were either in palaces or temples), and which therefore had no public presence (Robson 2013). Also, these libraries' collections were not similar like the ones in Roman libraries (Robson 2013), meaning that Roman libraries were open to the general public and Babylonian and Assyrian libraries were reserved only for the priests or government officials.

It must be mentioned that, although no catalogs or finding aids have been found as yet, many of the discovered tablets at Nineveh from the library of Ashurbanipal (reigned 668–627 BCE) were still in clay containers that had an attached clay tag recording a number; for multi-tablet works, the tag bore the number of the tablet, the first line of the previous tablet as well as the title of the entire work (Norris 1969). Tablets found at Kalhu (eighth-century BCE) had tags that recorded the names of the scribes that wrote the tablets as well as the from which they came (Robson 2013). At Edfu in Upper Egypt there is evidence of a catalog of books that is engraved on the library walls, but in all probability

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

this was just a list of the books (Norris 1969) rather than an attempt to organize the materials for retrieval.

Kim Ryholt sought to uncover the roots of the Alexandrian Library, built by the early Ptolemies in Egypt (323–222 BCE), with the long history of libraries in Egypt. Though established as a center of Greek learning, the Alexandrian Library certainly did share the long tradition that of libraries in ancient Egypt. However, the Egyptian libraries' intents resembled those of the Babylonian and Assyrian libraries: "On the contrary, there are indications that these libraries zealously protected their writings which were frequently described as 'secret' throughout the three millennia which our sources cover. The primary reason for restricting access to the literature may well have been to protect it from abuse and to retain its potency" (Ryholt 2013, 37).

Ryholt then notes the stories recounted by Strabo and Porphyry that Pythagoras, Plato and Eudoxus could not get the Egyptian priests to communicate any of their knowledge to them (Ryholt 2013). These priests, like those in Babylonia and Assyria, would not make public any information contained in the libraries. Because of this radically different role that libraries shared in Egyptian life, they would have had limited impact on scholarly work in regards to textural criticism, medicine, the sciences and lexicography, for which the later Alexandrian Library became famous and which had an impact on Greek scholarship (Ryholt 2013).

"In the Western world our patterns of thinking are derived from Greek philosophy and science, but this is not necessarily true of other cultures" (Broughton 2004, 16). The taxonomies created in the West are rooted in Aristotle's writings, which influenced Linnaeus and Sir Thomas Bacon. In the West, the Alexandrian Library catalog, the *Pinakes* of Callimachus, is considered the first library catalog.

One in a long line of librarians at the Alexandrian Library, Callimachus (ca. 305-ca. 240 BCE) wrote the *Pinakes* while a young man at court (Blum 1991). He also wrote a plethora of other works. Callimachus' *Pinakes* has been described as:

> ... this monumental, 120-vol. inventory of Greek literature categorized authors according to genre or discipline and arranged in alphabetical order. Each name was followed by a brief biography and a predominantly alphabetically arranged register of works. The incipit and number of lines were noted for each, as were questions of authenticity as necessary. Callimachus' critical inventory served simultaneously as a national 'bio-bibliography' and as an overview of the history of literature. (Cancik et al. 2006, vol. 6, col. 264)

Francis J. Witty classifies Callimachus as from "... a line of scholar-poets ... who were taught how to do scholarly research by Aristotle and his successors" (Witty

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

1973, 237). Classicists have focused solely on his poetic works, many of which are still extant, which cannot be said for his research works (Witty 1973, 237).

The Alexandrian Library had a large collection of scrolls on different subjects; there were also multiple copies of the same work. "Publishing" in the ancient sense meant having scribes, usually slaves, copy a work by hand. Authors would commission copies of their works to give to friends, libraries, or important people that they were trying to impress. Popular works were liberally copied as there was no concept of copyright in the modern sense; once an author released his work, anyone was free to make copies and, over time, discrepancies crept into the text, either accidentally or deliberately. The Library served the philosophers in the Museion, the "hall of the Muses," where these scholars tried to determine the identity of the original text and what had been added or changed over time. This is known as textural criticism, and multiple copies of a work were needed to do a comparative study in order to better determine the authentic lines.

The full name of the *Pinakes* was *Tables of Those Who Were Outstanding in Every Phase of Culture, and Their Writings—in 120 Books* (Witty 1958). The Alexandrian Library collection is believed to have been divided into the following subjects: epic writers; dramatic writers; writers on law; philosophical writers; historical writers; oratorical works; rhetorical works; and, miscellaneous works (Norris 1969).

There were further subdivisions under each subject, "... the entries were either arranged alphabetically under the names of the authors, or else, chronologically" (Norris 1969, 4). Witty lists an entry used by some scholars to "prove" that the Pinakes was an alphabetical catalog whose main entry was author. The bracketed references refer to the ancient writers who either quoted from or referred to the Pinakes; and where these references have been reprinted by Rudolf Pfiffer and given a number (Witty 1958):

> I know also that Callimachus in his *Pinax*
> *of Miscellaneous Literature* has included
> books on cake making by Aegimius, Hegesippus,
> Metrobius, and even Phaestus.
> [Athenaeus xiv. 643E; Pfeiffer 435] (Witty 1958, 135).

Callimachus recorded quite a bit of information about the author. None of the 120 scrolls have survived; parts of the *Pinakes* have been identified in parts of works written by other people. The ancients would "quote" a source without citing what, exactly, was the information that they were copying, as the above translated quote demonstrates. This has resulted in attempts to identify that which came from the *Pinakes* from the text of the author. Witty identifies and

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

reprints all the parts of Greek literature in which it is believed that the Pinakes is quoted or is being referenced. The following is example of an entry under the subject oratory:

> According to Callistratus, Diotimus claims
> that he (Demosthenes) delivered his first
> public oration before an audience of Athenians;
> and those who compiled the tables of
> orators entitled it *On the Symmories*.
> [Dionysius of Halicarnassus *Ad Ammaeum epist.* 4; Pfeiffer 432] (Witty 1958, 134).

Although Callimachus is not mentioned, it is believed that the tables being referred to were the *Pinakes* (Witty 1958). This can be seen here as well:

> Hecataeus of Miletus, in the *Description of*
> *Asia*–if the book was a genuine work of
> the historian, for Callimachus registers it
> under Nesiotes.
> [Athenaeus, Epitome of Book ii, 70B; Pfeiffer 437] (Witty 1958, 134).

Actually, this is also the only entry found that indicates that a "history" subject existed in the *Pinaks* (Witty 1958). The following has been attributed to the subject of philosophy:

> He (Parmenides) philosophizes in verse,
> just as Hesiod, Xenophanes, and Empedocles
> ... and it seems that he was the first to
> discover Hesperus and Phosphorus-so
> says Favorinus in the fifth book of his
> *Memorabilia;* but others say it was Pythagoras.
> Callimachus, however, says that the
> work was not his (Pythagoras).
> [Diogenes Laertius ix 22/23; Pfeiffer 442] (Witty 1958, 134).

According to Witty, of the 825 fragments that exist by or about Callimachus' work, only five are considered to contain *Pinakes* information (Witty 1958): "However, it is necessary at the very beginning to call attention to the fact that there are only about five true fragments, that is, actual quotations from the *Pinakes*. The other so called fragments are merely oblique references by ancient authors to the bibliographical work of Callimachus, which scholars have concluded to be references to the *Pinakes*" (Witty 1958, 132).

The above text from the *Pinakes* gives some indication of what was contained in an entry. The main entry of the work was entered under the name of the author, something which is true today, although this is changing in the web

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

environment. The title of the work is also an important access point, although secondary to the author's name. The subject area, where the book would have been placed on a shelf, was also indicated. Most probably the library collection was divided into subjects and then works were shelved alphabetically by the name of the author and then by title. Norris believes that in his entries Callimachus also summarized the work, provided a bibliography, and included some biographical information on the author (Norris 1969).

It was during Hellenistic times (323–30 BCE) that the idea of art as a commodity and something to beautify a public building or a home came into being, "by the mid-6th century BCE at the very latest, a sense of visual art as a piece of high culture infused the holy ground that yields our own record of 'Greek masterpieces' by 'great artists'" (Kuttner 2015, 52). Kuttner goes on to mention how the women in one of Herodas' third-century poems admire art in a temple (Kuttner 2015). Herodas' fourth mime does, indeed, depict two women, Phile and Kynno, in the Temple of Asklepios, come to make offerings while admiring the art:

> PHILE: Oh, what lovely statues, dear Kynno; what artist made this sculpture and who is the person who dedicated it?
> KYNNO: The sons of Praxiteles; don't you see the words on the base? And Euthies son of Prexon dedicated it ...
> KYNNO: See, Phile, that girl looking up at the apple: wouldn't you say that if she doesn't get the apple she will quickly expire?
> PHILE: And that old man, Kynno—
> KYNNO: Oh, by the Fates, how the child chokes the goose. Certainly if it were not stone before our feet, the work, you'd say, will speak. Ah, in time men will be able to put life even into stones. (Rusten, Cunningham, and Knox 1993, 257, 259)

And their admiration of the different statuary in the temple continues. The Temple of Asklepios was considered a museum, as famous statuary decorated the sanctuary. This is what Rome inherited from the Hellenistic rulers of the East who designated buildings where art was displayed for viewing, much like our modern museums.

The idea of art in libraries in part helped nineteenth-century archaeologists locate the Library of Pergamum behind the Temple of Athena. Here, there were four rooms, one of which was found what is believed to have been a copy of Athena Parthenos by Phidias, which stood in the Parthenon in Athens (Coqueugniot 2013). There was little questioning among scholars of this location for decades (Coqueugniot 2013). Because Athena and, later, Minerva, was goddess of wisdom, her statues stood in libraries, and the largest room, where the statue was believed to have stood (Coqueugniot 2013) was thought to be a reading and stacks room (Coqueugniot 2013):

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

> The side rooms were often considered as stack rooms for the library's collections, but it is equally possible to propose restitutions of these rooms as repositories for offerings, banquet rooms or 'annex rooms' of the sanctuary. Finally, there is no reason to consider that the existence of the *stoa* itself is linked with the existence of a library in its back rooms. (Coqueugniot 2013, 120)

Stoas, a columned, covered walkway or porch, in front of libraries or connected to peristyles allowed visitors to read by natural light. Most Greek and Roman libraries had one of these architectural features. The main room and the neighboring ones also had mosaic floors (Coqueugniot 2013). That this was the library's location is now doubted; the actual location of the library is unknown.

The ancient Greeks kept their books in gymnasia; however, there is no physical evidence from Hellenistic gymnasia that libraries were located in them (Petrain 2013). These gymnasia were shaped like Ls with stoas along the inside. It is known that, besides books, there was also art in the gymnasia. Catalogs, i. e. inventories, were created for these collections. There are two examples of this, one from Athens, ca. 170–135 BCE, and one from Delos, 156/155 BCE. In the Athenian example, statuary is standing in the gymnasium part known as the *exedra*, which is the area where there were seats, allowing people to converse. "*Exedrai* were the most conspicuous feature of a gymnasium; they were to its intellectual life what the *Balbis* was to the training of the body [*Balbis* was the running track]. It is precisely this part of the gymnasium that a wealthy benefactor to the educational life of a city chose for his gift of a statue ..." (Clay 1977, 263). There are other statues mentioned in the inventory, statues that were in both parts of the gymnasia, the *exedra* and the *balbis*: Artemis, Asklepios and Hygieia, and the Kouretes, who took care of the infant Zeus (Clay 1977).

In the Delos example, Athens had taken control of the island when an administrator had drawn up the inventory as part of a larger inventory of objects that were housed on the island (Clay 1977). Once again the gymnasium has statuary, this time bronze statues; the inventory also lists location, description, and donor names (Clay 1977). The difference between the Athenian statuary and the Delian in the respective gymnasia are that the Athenian statues were given to the city by donors whereas the Delian statuary is referred to as dedications, like in temples or shrines, and are dedications from the sculptors who created them (Clay 1977). Thus, the Delian figures have religious meaning whereas the Athenian statuary does not.

Nonetheless, why should there be exclusivity to religious art in these collections? Technically, Athenian statuary is of gods and other mythological themes as well as historic figures. These could be considered religious even though they were not dedicated as such. These statues were for decoration; they served no

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

other function but to beautify the gymnasia. This was art, on view to visitors, and the gymnasium was therefore in part an outdoor museum.

The Romans of the imperial era expected that art was supposed to be displayed for all to see, as the emperor Tiberius, who reigned from 42–37, discovered when he moved the *Apoxyomenos*, a statue created by Lysippus, to his private residence (Macaulay-Lewis 2009). The Romans created their own buildings specifically designed to hold libraries, thereby breaking with the tradition of the Greeks (Petrain 2013). "Roman public libraries therefore seem to have been designed from the outset with an ostentatious public function in mind, as monuments displaying works of art and providing a magnificent backdrop for cultural activities of various sorts" (Nicholls 2013, 264). In addition to the statuary, there were portraits hung around the library, perhaps made of gold or silver or even bronze, of the authors whose works were kept in wooden cabinets in niches around the room; everyone therefore had an idealized idea of how the authors looked (Petrain 2013). Indeed, in the section on portraiture in *Natural History*, Pliny discusses the library in Asinius Pollio's atrium (Petrain 2013): "For Pliny, the visual experience provided by the library is as important as the literary monuments to be found there, the novel *imagines* of authors as stirring as those of the ancestors once were ..." (339). David Petrain (2013) suggests that the portraits of these authors may have created a "...'metonymic habit', a readiness, even a keenly felt need to perceive close connections between literary works and the visual supplements associated with them" (340).

The display of art and the beauty of the surroundings was a major role of these libraries, beyond the collection and access to materials. Matthew Nicholls (2013) points out that if the Romans had wanted to protect their library collections from theft and water damage, they could have located them inside buildings and on upper floors, but then they could not be seen through the grand space accessible from street level (Nicholls 2013). Perhaps just as importantly, the Roman emperors, in many cases the benefactors of libraries throughout the empire, were experts at using art for political effect: "Rome's public libraries formed part of a grand backdrop suitable for the conduct of public business, setting a pattern for their successors" (Nicholls 2013, 267).

Those Roman public libraries, like the Alexandrian Library, were organized by general subjects, then subdivided by author. However, unlike in the Alexandrian Library, the subjects by which the Roman libraries were organized are not known (Harris 1995). When Herculaneum's Villa of the Papyri's library was excavated, thin metal plates were found which is believed to have been library's shelf labels, as well as other indications of organization. "Unfortunately, inadequate documentation when the papyri were removed has destroyed the possibility of collocation studies" (Sider 1990, 538). One major difference between

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

Greek and Roman libraries is that Roman libraries were divided by language: one area for Greek, the other area for Latin, under which were the subject divisions and author subdivisions; all works by one author were kept together and the works on the schools of philosophy were kept together (Harris 1995). As for catalogs: "Catalogs of two types were known, and sometimes both were used. The first type was a sort of classified catalog, or shelf list, arranged just as the rolls themselves were stored. The other was a bibliographical catalog, arranged by author but giving titles or first lines. lengths of works, and sometimes bibliographical information about the author" (Harris 1995, 65–66).

As for ancient archives, some of the oldest records originate from the Near East, where techniques in archival management developed and were later transmitted to Persia, Egypt, and Mycenaean Greece, and well into the Hellenistic Age (Brosius 2003). Of the 253 archives and libraries in the Near East that are discussed by Olof Pedersén, 127 are believed to be from official government buildings and not private dwellings (Pedersen 1998). Unfortunately, not much is known about them:

> Although extant archival materials are extensive and reflect a range of ancient societies, modern scholars know surprisingly little about the purpose, functioning, and management of these archives. The number of unanswered questions far exceeds those for which we can marshal a convincing explanation. Few modern scholars have addressed the issues of archival traditions rigorously and in detail, or have dealt with questions of continuity across a comparable range of time, space, and political systems. (Brosius 2003, 1)

This is changing. Ebla, in Syria, was one of the first archives (ca. 2600-ca. 2240 BCE). This archive centers on the distribution and delivery of textiles and metals (Archi 2003). Ebla competed with Mari, an empire centered to the north, and would periodically acknowledge the suzerainty of its northern neighbor (Archi 2003).

The Ebla tablets can be dated from the names of ministers listed (Archi 2003). Alfonso Archi also believes that the shelf arrangement helped archivists to determine the dates of the tablets (Archi 2003). One problem is that the wooden shelves were destroyed, scattering the tablets on the floor and thus destroying the original arrangement (Archi 2003). What is most interesting is the shapes of the tablets used in the differing transactions. At Ebla, four different tablet shapes in the archive were used for differing transactions, which was a way to distinguish by sight what kind of transaction that was represented: textile distribution, distribution of precious metals, income accounts; and sheep deliveries (Archi 2003). Additionally, personal archival tablets from the Neo-Babylonian period had certain phrases or words, formulae or clauses (Baker 2003), that indicated the transaction.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

The most famous archive of the ancient Greeks is the Metroon (300s BCE-267) in Athens. The Metroon was sacred to "the mother of the gods," and held a cult statue of the mother and an altar. The other rooms were used for storing the official records of the Boule, one of the elected bodies of the Athenian government.

It should be noted that there was a major difference between the ancient Near Eastern and Greek archives. Whereas Greek archives predominately housed decrees and other official documents, Near Eastern archives were predominately economic in focus: "The classical world tends to apply the term 'archives' to collections of legal documents and decrees. By contrast, ancient Near Eastern archives were concerned with the documentation, processing, and storage of predominantly economic texts. These include letter orders and receipts for goods, livestock, tax payments, and payments for labourers" (Brosius 2003, 5). The Athenians were keeping records, as the Metroon attests, but this was not the only archive in Athens. There were many archives in Athens as the magistrates and other legal bodies kept their own archival collections:

> Decentralization does not mean that individual collections of laws were themselves poorly organized; we have no evidence of how magistrates maintained copies of laws in their possession, so it is safe to attribute to them neither too high nor low a degree of order. But decentralization could have created uncertainties about what laws were valid or where the texts of particular laws could be found, and problems arising from such difficulties may have contributed to the Athenians' decision to review their laws in the last decade of the fifth century. (Sickinger 1999, 84)

What the Metroon kept, after the reforms of Cleisthenes, were the documents generated by the Boule, the Athenian city council, and the Ekklesia, the Assembly (Sickinger 1999). The question is how these documents were arranged. "On this issue, scholars have reached negative conclusions" (Sickinger 1999, 83–84). Nonetheless, dating these documents is not impossible:

> In the fifth century, decree prescripts generally included the name of the secretary of the Boule in office when the decree was enacted, the name of the tribe that held the prytany at the same time, and the name of the chairman (epistates) of the prytaneis who presided at the meeting of the Ekklesia at which the decree was ratified. The order in which these items appeared shows some variation, and occasionally one or more items are left out. But each of them is generally present, and since it is generally agreed that the prescripts were added to the texts of the decrees by the secretary of the Boule himself, their contents may offer some insights into the records from which published decrees were copied. (Sickinger 1999, 84)

This suggests some type of archival organization, allowing for the retrieval of information.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

Over time, the Metroon came to hold private documents such as wills. That an official government archive would hold documents of citizens is not as strange as it may seem to our modern understanding of an archive. "The phenomenon of private documents deposited in public spaces, and, more surprisingly, vice versa, seems to be a practice which can be found throughout the history of record-keeping" (Brosius 2003, 11).

China's *Seven Epitomes* and Archives are described as a type of classification; "Classification is a map of human knowledge which provides a panoramic view of man's thought and experience and in which the intellectual achievements and tendencies of a given period or region will be reflected. Its history, consequently, runs parallel with the development of human thought, in spite of geographical and cultural differences" (Tsien 1952, 307). China had a history of organizing information as long as that of the Near East. The Warring States Period (475–221 BCE) had given rise to a plethora of philosophical schools, including the one headed by Confucius (551–479 BCE) (Jiang 2007).

Several dynasties had ruled China, but the Chinese were finally unified in 221 BCE by the Qin Dynasty, which standardized Chinese characters and used stone inscriptions to record events (Zhang 2004). Unfortunately, the Qin were a ruthless dictatorship that ordered the destruction of all non-Qin books, including those in private hands, and only allowed medical, agricultural and divination books to circulate (Zhang 2004). The Qin were overthrown and the Han Dynasty eventually came to power. In 140 BCE the Han Emperor Wu made Confucian texts the cornerstone of learning for the civil service, thus transforming Confucianism into becoming the official orthodox philosophy of China (Jiang 2007).

Chinese tradition holds that during the original Three Dynasties of Xia (twentieth-sixteenth centuries BCE), China was ruled by sage-kings who created a peaceful, unified society that was completely harmonized; there was no dissention or disorder, only unity and Dao, the Way (Jiang 2007). Decline led to disorder and unhappiness, and in an attempt to restore the Dao and bring harmony back to society as it had been during the golden age, Confucius worked to preserve whatever extant works survived from this earlier time; also an archivist, Confucius edited works based on the archives of previous dynasties (Zhang 2004). These books, the *Odes*, the *Documents*, the *Rites*, the *Changes*, and the *Spring and Autumn Annals*, became known as the Five Classics during the Han period, and were believed to hold the truths that would lead back to the perfect society:

> Turning to the past to look for answers to present issues was not just a practical solution but also a cardinal concept of Confucianism. In the broader Chinese worldview—Confu-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

> cianism and Taoism alike—the universe originated from the ultimate oneness, the Dao. For either a society or an ideology, the decentralization of power or the pluralism of philosophy was a sign of the decline of the Dao. (Jiang 2007, 13)

The Qin had destroyed much of the older works before they were overthrown. The Han therefore used Confucian principles to guide government and establish a civil service system to recruit scholars into the imperial bureaucracy (Zhang 2004). The chosen holdings in the imperial library would become the defining texts for study for the civil service.

The *Seven Epitomes* was a working government document and it was the catalog to the Han imperial library created at the request of the emperor (Lee 2010). A parallel between the structure of government agencies, real and imaginary, and the structure of texts was created because writing had been linked for centuries with government agencies, which evolved over time and expanded with government, before private individuals began their own activities (Lee 2010). To classicists like Liu Xin, any documents created during the perceived idealized government were still relevant and therefore needed to be preserved (Lee 2010).

Confucianism was the guiding principle behind the creation of the catalog (Lee and Lan 2009). The Han wanted to preserve and restore those writings that were lost before the dynasty came to power. "The establishment of imperial collections and national bibliographies was an effort by the Han imperial government to restore and preserve the written culture from the past, particularly the period extending from Confucius (551–479 B.C.) to the end of the Warring States Period (221 B.C.)" (Jiang 2007, 2). Emperor Cheng appointed Liu Xiang to lead scholars in compiling a review of books in the imperial library; Liu Xin, Liu's son, finished what became known as the Seven Epitomes (Jiang 2007). In reality, there were only six epitomes, "... the very first epitome is said not to represent a class. In fact, a popular belief by bibliographers is that the first epitome (the Collective Epitome) was a collection of introductions to the entire catalog and to its individual classes and divisions" (Lee 2010, 202). The epitomes are listed with the number of divisions (Lee 2010): the Six Arts (nine), the Masters (ten), Lyrics and Rhapsodies (five), Military Texts (four), Divination and Numbers (six) and, Formulas and Techniques (four).

Hur-Li Lee (2010) analyzed the Epitome of Divination. The divisions in the class are "from Heaven to Earth," (202) which is why astrology and chronology came at the beginning and geomancy was at the end. Divination goes far back into Chinese history and served a variety of functions:

> At court, divination was not only a way to give advice for mundane activities but also a strategy for legitimizing and perpetuating the ruling of the royal house. As such, diviners

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

> were able to hold high positions and tremendous power in the government ... Because divination had been central to the state for hundreds of years and divination manuals existed in large number ... the important of divination in the text culture of the Han was unquestionable. (Lee 2010, 205)

There were six divisions under the Epitome of Divination:

- Patterns of Heaven (including astrology, astronomy, meteoromancy)
- Chronology (including calendars, astrological and hemerological calculations, related mathematics, and genealogical tables of rulers);
- Five Phases (including various numerological calculation modeling for the two divisions above);
- Milfoil and Turtle Shell (divinations using milfoil stalks and turtle shells);
- Diverse Prognostications (including demonology, oneiromancy, incantation, and exorcism);
- System of forms (including sitting and physiognomy). (Lee 2010, 204–205)

There were divination works classed into epitomes other than the one for divination. For example the Book of Changes, though a theoretical work about divination, was considered one of the Six Classics and was therefore classified into the Epitome of the Six Arts (Lee 2010). "Mathematics was part of the Epitome of Divination only as it was applied to astronomy and harmonics. Nowhere else in the Seven Epitomes did mathematics hold a spot.... The departmentalization of the state machine was a significant consideration in the cataloging project. The influence of government was apparent in the classification of the three types of technical writings and in the exclusion of others" (Lee 2010, 208).

During the Jin Dynasty (265–317), the *Seven Epitomes* was modified into four divisions by Xun Xu in his *New List of the Central Collection of Classics* (*Zhongjing xinbu*) (Jiang 2007, 7). Xun eliminated the names of the classes, as he moved around some of the subjects: one (six scriptures), two (various masters, military writings, divination and calculation, medicine and formulas), three (history [from six scriptures]) and, four (odes).

Emperor Mu later commanded Li Chong to create a bibliography of the imperial library back to the time of Emperor Yuan; thus, the *Bibliography of Emperor Yuan's Library* (*Yuan di shumu*) was created. This time, Li reorganized the four divisions: one (remains as one), two (becomes three), three (becomes two), four (remains four) (Jiang 2007). Even later, the *Jing ji zhi* would give names to Li's divisions: one becomes Classics, two becomes History, three becomes Masters, four becomes Collections of Writings (Jiang 2007).

Although called national bibliographies which gave birth to the *Seven Epitomes* and later its four-part successor (Jiang 2007), in reality these bibliographies, which were catalogs, were of the materials in the imperial library that

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

supported the Confucian texts and the idea of Dao; this was not a national bibliography in the sense that it was a comprehensive list of the materials were of everything written in Chinese. The *Seven Epitomes* only indexed materials in the imperial library.

On the other hand, the *Pinakes*, Callimachus' catalog, was a national bibliography in the sense that the Alexandrian Library attempted to bring together the entire corpus of Greek writing back to the time of Homer; the *Pinakes* was the catalog of the materials that were in the library. There was no Greek nation, as the Greeks were divided into city-states and the successor states of Alexander's empire. However, Greece as a culture, language, and religion, was reflected in the *Pinakes*, thereby making the catalog *a defacto* national bibliography.

Most experts agree that the "Bibliographic Treatise," (Lee 2010, 202) part of the *History of the Former Han Dynasty*, is an abridgement of the *Seven Epitomes*. "The original classificatory structure, its six classes and divisions and the vast majority of the entries, although shortened, remained intact" (Lee 2010, 202). Ironically, it was just this simplification that spelled the end of the *Seven Epitomes*, disappearing sometime between the late ninth and the first half of the tenth century:

> The *Seven Epitomes* was one of the most influential written records in imperial China. Its importance, however, has not always been apparent or fully appreciated.... Few made an effort to understand its design fundamentals. As an unfortunate result, its abridged version, the Han Treatise, took precedence and later made the catalog seem dispensable, leading to its eventual disappearance almost a thousand years ago. (Lee 2010, 208)

In some ways, the *Seven Epitomes'* classification scheme parallels that of the United States Government's Superintendent of Documents (SuDoc) classification system. Both systems organized texts by government office, "... demonstrating the commonality of institutional consideration in classification across some cultures" (Lee 2010, 209).

The Chinese were as advanced in archival organization as they were in libraries. During the Western Zhou Dynasty (eleventh century-771 BCE), bronze objects were used to record archival inscriptions. "The bronze ware had evolved into a cherished recording medium with lengthy writings covering a wide range of subjects, including among others: proclamation, bestowal, litigation, military expedition and adventure" (Zhang 2004, 19). It should be noted that this bronze ware was art that was being used as both decoration and for a practical purpose. While other documents would be destroyed, the bronze ware was usually carried off to decorate a conquering lord's castle or public area, thus continuing the bronze's function as art. Official events were recorded on these customized

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

vessels (Zhang 2004). Besides bronze ware, bones were used for oracular writings, while wood and bamboo strips and, later, paper, were used in keeping records (Zhang 2004). It also became policy to make copies of official documents, keeping the originals in the central archive and filing the duplicates in different administrative offices (Zhang 2004).

The Tang Dynasty (618–907) further introduced new ideas into archival management. Different color-coded papers were used by different ministries, and only a single issue could be addressed in one document (Zhang 2004). "The filing system required that a document in a depository be dated, signed, stamped, and filed monthly, and the total account reported to the central agency at the end of the year. Imperial decrees, memos to the throne, official appointments, financial records, marriage registrations, land accounts, and legal cases were deemed permanent records; everything else would be weeded out every three years" (Zhang 2004, 27). There were other collections of records, such as taxation, and map collections as well as records arranged by feudal rank and family location while military records were held by the Ministry of Defense (Zhang 2004).

The Sung Dynasty (960–1279) in 1006 established a central archive at the Golden Gate of the palace in the first capital, Dong Jing/Kai Feng (Zhang 2004). Archivists still came from family inheritance, but there was an annual recruitment by the government that included both oral and written tests (Zhang 2004). It was the Song who changed archival form from rolls to paper. Types of paper, depending on the rank and office, were further subdivided into five types and twelve classes; the imperial palace had seven types of categories and twelve classes (Zhang 2004). Perhaps most amazing of all, the Song, unlike former dynasties, did not keep their archives in closed wood or metal containers but in open stacks like modern archives (Zhang 2004). Multi-levels of stacks were used as well as a comprehensive filing and indexing system that was developed (Zhang 2004). And just like the Tang, the Song archives contained official appointments, land and taxation records, marriage and lawsuit cases (Zhang 2004).

Of catalogs and dictionary catalogs

In Europe, it was the medieval monasteries that would create catalogs for their books. The catalogs would evolve over time from simple inventories into actual catalogs that were meant to help in the location of the books. Monasteries did not have separate rooms for books until the fourteenth century; prior to that,

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

books were kept in cupboards around the church and the cloisters (Norris 1969). It was only when holdings of the library became large that books were classified by subject:

> At first the division might be between theological works and secular ones; between Latin works and those in other languages; or between textbooks and more serious tomes. The religious works might be subdivided into Scriptures, commentaries, biographies, and service books. The secular works, particularly in the cathedral libraries, might be divided according to the teaching subjects of the trivium and the quadrivium (grammar rhetoric, and dialectic; arithmetic, geometry, music, and astronomy). In some libraries, these divisions were designated by letters and these letters were prominently inscribed on the book-chests. Such a broad classification system apparently sufficed as long as the collections were small, but later in the Middle Ages more complicated schemes were planned, if not actually used. (Harris 1995, 100–101)

The trivium was the basis for classical education, as it was studied in ancient Greece. The quadrivium is the next level of education after the trivium. Together they formed the seven liberal arts. The purpose of the study of these subjects was to create an educated person who could think.

The entries in the catalog of Christchurch, Canterbury, created sometime between 1313 and 1331, were in alphabetical order, a unique way to list materials for medieval times (Norris 1969). The catalog was divided into two Demonstrations, which might indicate two different areas where the books were kept; the Demonstrations were divided into Distinctions which may have been book shelves or subject classes (Norris 1969). There are several ways that were used to identify a book that was being listed:

> Instead of the second folio words, the opening words of the book are given; these words are introduced by various phrases, 'qui sic incipit' [He begins] is used when the work consists only of one volume; 'quorum unum ... et aliud' [one ... and other] is used when the work consists of more than one volume. In many instances, the opening words are enclosed in inverted commas without being preceded by any introductory phrase, and in some cases the opening words have been omitted by mistake. The only other bibliographical details given are the size, and the number of volumes of which a work consists. (Norris 1969, 46)

In many cases, there were no titles for the books, prompting descriptions of what the book looked like or what it contained. One major problem was composite books, or books with more than one title bound within them. In some cases, the cataloger gave no indication that there were other titles in the binding. The 1372 catalog of the Austin Friars solved this problem by using multiple letters to mark the catalog with composite works in one volume (Norris 1969).

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

The subject divisions in the 1391 Durham Cathedral library catalog were written in the books as well as the opening lines of second folios, or pages, being listed in the catalog to differentiate between different copies of the same work (Norris 1969). The Benedictines, who founded Durham, also allowed the borrowing of materials, and the notes of those borrowings are in the catalogs, designated by *Accomodatur*, *habet*, *Est vetus liber*, and *Ponitur in libraria* (Norris 1969). This is in line with the Benedictine belief in loaning books to people outside of the monastery (Norris 1969).

Subject catalogs were in favor in the fourteen century (Norris 1969). Alphabetical order is not the default under subjects. Only once does alphabetical order appear, and this is in the Christchurch catalog under the subject of Theology (Norris 1969). The *Registrum Librorum Angliae* was probably the creation of the Franciscans who took 183 libraries in England, divided them into eight groups called custodia, and then created a catalog of works by authors, listing their titles and which libraries had the work. All libraries were assigned a unique number (Norris 1969). Normally the first line of the second page of a work was listed as a way to differentiate any second copies, but this is missing from the catalog, since its purpose was to inform where copies of works were located (Norris 1969).

The *Tabula Septem Custodiarum Super Bibliam* is considered a later work than the *Registrum* if only because it has an alphabetical order, something which is unseen at the time of the *Registrum* (Norris 1969). Nonetheless, its function was the same as that of the *Registrum*. The title is misnamed and there are eight custodies listed (Norris 1969). Because copies and not the original manuscripts are still extant, it is hard to tell if the *Registrum* and the *Tabula* were ever completed. Many titles have no library numbers beside them, and many libraries with numbers have no listings (Norris 1969). The *Catalogus Scriptorum Ecclesiae*, written by John Boston of Bury, who lived around 1410, was a continuation of the *Registrum* (Norris 1969). Boston was a Benedictine, so the custodies were eliminated, but he kept the numbering system of the *Registrum* (Norris 1969). Boston omitted some entries from the *Registrum* and the *Tabula* while adding new ones (Norris 1969). The *Registrum*, *Tabula*, and *Catalogus* are early examples of union catalogs, or a catalog with the holdings of several libraries listed. These types of catalogs would be very helpful to users because materials needed by a user could be easily located at neighboring libraries. Printed union catalogs were so useful that they would last well into the twentieth century.

Conrad Gesner (1516–1565) was a botanist and also a bibliographer who wrote the *Bibliotheca Universalis* (1545) and its supplement, the *Pandectarium* (1548–1549), which was a catalog of all Greek and Latin authors and their works

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

(Norris 1969). Gesner's work was much different than the *Tabula*, the *Registrum*, and the *Catalogus*. His idea was to have libraries duplicate his catalog and add their own marks to items that they owned (Norris 1969). The catalog first divided a library collection by size: M, or magnus, and P, or parvus, with the smaller collection coming after; then recorded the order of the books on the shelves, after which was an alphabetical index by author's name; and finally an accessions list of when the items were added (Norris 1969). There were problems with this arrangement, one being that when M books were added, the shelving numbers of all the P books changed (Norris 1969).

With the advent of the Renaissance, Reformation, and Age of Reason, the rediscovery of Greece and Rome as well as a renewed interest in science, started a process that increased interest in scientific classification and, by extension, library classification. Sir Francis Bacon (1561–1626) and Carl Linnaeus (1707–1778) rediscovered Aristotle's writings, which influenced their own works on scientific classification as well as others such as Gesner. Increased literacy meant that more and more books were being published, thanks to the printing press, thereby increasing the number of works in libraries. The time of libraries with only a few hundred books was past. Newer, more robust systems of classification were needed to handle the increase in volume.

Bacon, in particular, had a large influence on early library classification. The Pennsylvania Library Company attributed their classification to Bacon's scientific divisions (Norris 1969) and Yale looked to Bacon in its creation of subjects (Norris 1969). Thomas Jefferson was also influenced by Bacon. The 1783 catalog of his personal library shows classification influence from Bacon (Servies 1950). Jefferson was familiar with Bacon's classification on human facilities at least as early as 1781, as is evidenced from his writings (Servies 1950). The Library of Congress (LC) collection, since at least 1808, was originally organized alphabetically by size (Servies 1950). With the burning of the library by the British in 1812, Jefferson sold his collection to the United States government to replace the one that had been destroyed. The eighteen subject classes show a strong Baconian influence. It would be Jefferson's classifications, with modifications, that would be used by LC up through the nineteenth century, only being abandoned with the creation of the Library of Congress Control Number (LCCN).

In 1791, the French government issued the first national cataloging code; "the code provided guidelines on how to describe publications, instructions on how to mark shelves to facilitate the location of a particular book, and a brief statement on the qualifications desirable for catalogers" (Hopkins 1992, 378). No longer would individual libraries have to create their own organizational system and record what they thought might be important. The code was also

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

responsible for the creation of the card catalog; playing cards were used to record the information (Hopkins 1992). The card catalog would soon be adopted by libraries as the way to easily update their holdings. It would also allow libraries to file cards under multiple authors, titles and partial titles, a number of subjects, and series titles. In the future, journals, films, microforms, and all types of materials would be cataloged in this way.

However, there were two major drawbacks to the card catalog. It was very labor-intensive: people needed to know how to correctly file cards in the catalog. Dashes, periods, semicolons, and other grammatical marks all meant different things in the cataloging order, and a person had to be trained on how to do it correctly. It was also impossible for people to file cards for their whole work day. Fatigue and eye strain limited the filing to be done usually early in the morning when the library first opened and only for a few hours. The second drawback was that, like the printed catalogs of old, the card catalog was also out of date. One title could generate several numbers of cards that needed to be filed in several drawers. Unlike materials that were added in the old medieval libraries, materials are added in the hundreds every week in the big university and research libraries. With full card sets arriving daily, the card catalog never was caught up to what was in the collection. Still, the resilience of the card catalog was admirable. It was only on October 1, 2015, after over 150 years, that the Online Computer Library Center (OCLC) printed the last cards for card catalogs (Coyle 2016).

Dictionary catalogs were still useful, and would remain so until nearly the end of the twentieth century. The only way a library could let other libraries know of their holdings was through the printing of dictionary catalogs. For university and research libraries, these dictionary catalogs would run into the hundreds of volumes. In 1979, the New York Public Library (NYPL) printed *The Dictionary Catalog of the Research Libraries of the New York Public Library, 1911–1971*. The total number of volumes was 800. It should be noted that the holdings of the branch libraries of NYPL were not included in the catalog. These 37 x 26 cm. (13 ¼ x 10 ½ in.) volumes were made up of photocopied cards from the card catalog. Printed OCLC catalog cards had a date of production at the bottom, which made pulling certain cards for photocopying easy. Drawers of catalog cards would be removed on occasion with a note that the drawers would not be accessible while the cards were being photocopied.

In the United States, there were several national dictionary catalogs, such the *National Union Catalog* (NUC) and the *National Union Catalog of Manuscript Collections* (NUCMC) that were compiled by the Library of Congress with the help of the American Library Association. These voluminous works included the holdings of the Library of Congress and other large, academic libraries. Cop-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

ies were sold to libraries throughout the United States and abroad. The NUC broke up into different parts and the last NUC update was in 1983. The last NUCMC update was in 1993. Because these catalogs were issued in different parts, knowing where to search and what to search was imperative, which meant there was some training necessary in order to use the catalog. Also, there were symbols used for the different libraries, requiring an index to interpret them.

Conclusion

> I will begin with a simple, if bold statement that knowledge organization (KO) is pervasive. We know that KO is pervasive in human experience, as every act of cognition holds the possibility of categorization, even if only into yes or no 'sets' of perceived observations. But I believe we can go so far as to say that knowledge organization is as pervasive as knowledge itself.... The purpose of making such an assertion here is to remind us that classical knowledge organization systems (KOSs) extend far beyond the realm of bibliographic classifications, which consume the majority of scholarly energy in the knowledge organization community, are not even the proverbial tip of the iceberg. (Smiraglia 2016, xxi)

There is a long history of catalogs being used to order knowledge, from the Ebla archives, the *Pinakes* and the *Seven Epitomes*, and the catalogs of medieval times and the dictionary catalogs of the last century. Libraries, archives, and museums did use catalogs to organize information. Although discussing a folk classification, Vanda Broughton illustrates how what we see and what others of the past saw is not necessarily the same:

> Sometimes the basis of a folk classification completely escapes the scientific mind, as in the case of this classification belonging to the Dyirbal people of Australia. It combines a folk taxonomy with a linguistic classification, as is common in primitive cultures, and provides a spectacular example of an alternative view of the world. The classification contains four classes:
> – men, kangaroos, possums, bats, most snakes, most fish, some birds, most insects, the moon, storms, rainbows, boomerangs;
> – women, bandicoots, dogs, platypuses, echidnas, some snakes, some fish, most birds, fireflies, scorpions, crickets, the hairy mary grub, anything connected with water or fire, the sun and stars;
> – all edible fruit and the plants that bear them, ferns, honey, cigarettes, wine, cake;
> – parts of the body, meat, bees, wind, yamsticks, some spears, most trees, grass, mud, stones, noises, language. (Broughton 2004, 16–17)

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

Broughton then tries to make sense of the Dyirbal worldview:

> The first group evidently contains mainly boys' stuff. The second group looks generally rather more dangerous and unreliable than the first (although it is hard to see why 'most birds' are here), and group three consists of all good things. What the members of group four have in common is a mystery. But our inability to see the logic in this classification doesn't invalidate it; the principles on which it's based are just not known to us. (Broughton 2004, 17)

If four categories are kept and these items are organized into a Western pattern, the pattern might be:

- men, women, kangaroos, possums, birds, bandicoots, dogs, platypuses, echidnas, insects, fireflies, scorpions, crickets, bees, the hairy mary grub, bats, snakes, fish [living creatures];
- wind, water or fire, the sun and stars the moon, storms, rainbows, grass, mud, stones, noises, all edible fruit and the plants that bear them, ferns, most trees [the natural world]
- parts of the body, meat, yam sticks, spears, language, boomerangs; honey, cigarettes, wine, cake [miscellanea]

A more specific breakdown, eliminating the four categories, would be:

- men, women, dogs, platypuses, kangaroos, possums, bats, bandicoots, echidnas [animals that are mammals];
- snakes, fish, birds, scorpions [animals that are not mammals];
- insects, crickets, the hairy mary grub, fireflies, bees [animals that are insects];
- all edible fruit and the plants that bear them, ferns, trees, grass [items that are plants];
- storms, wind, rainbows, water, fire, mud, stones [things that are part of the natural environment];
- the moon, the sun and stars [celestial objects]
- yamsticks, spears, boomerangs, cigarettes [things made by humans];
- honey, wine, cake, meat [things people eat];
- noises, language [sounds];
- parts of the body [anatomy].

Not all of the items fit neatly into a Western breakdown, and some categories, such as the divisions "some insects" and "most insects," make no sense from the Western perspective. However, to the Dyirbal people of Australia, their folk taxonomy makes sense to them, just as the two examples make sense to the Western mind. It can be argued that the Western, scientific divisions are, in

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

themselves, folk classifications or folk taxonomies. Children are taught at a young age how to group like items together, which is no different than what the Dyirbal people teach their children to do. To the Dyirbal people, there is an inherent logic to their classifications, a logic we might not understand. This does not make their view of the world wrong. Scientific taxonomy may claim that it is detached from social knowledge, but this is not true. When first created scientific taxonomy was detached, but it is now common knowledge in the West, so common that children are automatically indoctrinated into it.

Organizing information can take many forms. However, the organization must make sense to the people who will use the information to find what they need; it must be user-friendly enough for users to locate the materials. Also, just because something appears to have no organization does not mean that it does not have one.

References

Archi, Alfonso. 2003. "Archival Record-Keeping at Ebla 2400–2350 BC." In *Ancient Archives and Archival Traditions: Concepts of Record-Keeping in the Ancient World*, edited by Maria Brosius, 17–36. Oxford: Oxford University Press.

Baker, Heather D. 2003. "Record-Keeping Practices as Revealed by the Neo-Babylonian Private Archival Documents." In *Ancient Archives and Archival Traditions: Concepts of Record-Keeping in the Ancient World*, edited by Maria Brosius, 241–63. Oxford: Oxford University Press.

Blum, Rudolf. 1991. *Kallimachos: The Alexandrian Library and the Origins of Bibliography*. Translated by Hans H. Wellisch. Madison: University of Wisconsin Press.

Brosius, Maria. 2003. "Ancient Archvies and Concepts of Record-Keeping: An Introduction." In *Ancient Archives and Archival Traditions: Concepts of Record-Keeping in the Ancient World*, edited by Maria Brosius, 1–16. Oxford: Oxford University Press.

Broughton, Vanda. 2004. *Essential Classification*. New York: Neal-Schuman.

Cancik, Hubert, Helmuth Schneider, Christine F. Salazar, and David E. Orton. 2006. *Brill's New Pauly: Encyclopaedia of the Ancient World. Antiquity*. Vol. 6. Boston: Brill.

Clarke, Rachel Ivy. 2014. "Find, Identify, Select … Socialize?: Alternative Objectives of Library Catalogs." *Proceedings of the American Society for Information Science and Technology* 51 (1): 1–5. doi:10.1002/meet.2014.14505101076.

Clay, Discin. 1977. "A Gymnasium Inventory from the Athenian Agora." *Hesperia* 46 (3): 259–67.

Coqueugniot, Gaelle. 2013. "Where Was the Royal Library of Pergamumo? : An Institution Found and Lost Again." In *Ancient Libraries*, edited by Jason König, Aikaterini Oikonomopoulou, and Greg Woolf, 109–23. Cambridge: Cambridge University Press.

Cotton des Houssayes, Jean-Baptiste Kent. 1906. *The Duties and Qualifications of a Librarian: A Discourse Pronounced in the General Assembly of the Sorbonne December 23, 1780*.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

Translated by Pierre-Alexandre Gratet-Duplessis. Literature of Libraries in the Seventeenth and Eighteenth Centuries. Chicago: A.C. McClurg & Co.

Coyle, Karen. 2016. "The Evolving Catalog: Cataloging Tech from Scrolls to Computers." *American Libraries*, February, 48–53.

Harris, Michael H. 1995. *History of Libraries in the Western World*. London: Scarecrow Press.

Hopkins, Judith. 1992. "The 1791 French Cataloging Code and the Origins of the Card Catalog." *Libraries & Culture* 27 (4): 378–404.

Jiang, Shuyong. 2007. "Into the Source and History of Chinese Culture: Knowledge Classification in Ancient China." *Libraries & the Cultural Record* 42 (1): 1–20.

Kuttner, Ann. 2015. "Hellenistic Court Collecting from Alexandros to the Attalids." In *Museum Archetypes and Collecting in the Ancient World*, edited by Maia W. Gahtan and Donatella Pegazzano, 45–53. Monumenta Graeca et Romana, Vol. 21. Netherlands: Brill.

Lee, Hur-Li. 2010. "Divination and the State: Classifying Technical Texts in Han China." *Library Resources and Technical Services* 54 (4): 200–211.

Lee, Hur-Li, and Wen-Chin Lan. 2009. "Purposes and Bibliographic Objectives of a Pioneer Library Catalog in China." *The Library Quarterly: Information, Community, Policy* 79 (2): 205–31.

Macaulay-Lewis, Elizabeth. 2009. "Political Museums: Porticos, Gardens and the Public Display of Art in Ancient Rome." In *Collecting and Dynastic Ambition*, edited by Susan Bracken, Andrea Gáldy, and Adriana Turpin, 1–21. Newcastle upon Tyne: Cambridge Scholars.

Nicholls, Matthew. 2013. "Roman Libraries as Public Buildings in the Cities of the Empire." In *Ancient Libraries*, edited by Jason König, Aikaterini Oikonomopoulou, and Greg Woolf, 261–76. Cambridge; New York: Cambridge University Press.

Norris, Dorothy May. 1969. *A History of Cataloging and Cataloging Methods, 1100–1850: With an Introductory Survey of Ancient Times*. London: Grafton & Co.

Nowak, Maria J. 1997. "The History of the Jagiellonian Library." *Libraries & Culture* 32 (1): 94–106.

Pedersen, Olof. 1998. *Archives and Libraries in the Ancient Near East, 1500–300 B.C.* Bethesda: CDL Press.

Petrain, David. 2013. "Visual Supplementation and Metonymy in the Roman Public Library." In *Ancient Libraries*, edited by Jason König, Aikaterini Oikonomopoulou, and Greg Woolf, 332–46. Cambridge: Cambridge University Press.

Robson, Eleanor. 2013. "Reading the Libraries of Assyria and Babylonia." In *Ancient Libraries*, edited by Jason König, Aikaterini Oikonomopoulou, and Greg Woolf, 38–56. Cambridge: Cambridge University Press.

Rusten, Jeffrey, I. C. Cunningham, and A. D. Knox, trans. 1993. *Characters / Theophrastus. Mimes / Herodas. Cercidas and the Choliambic poets*. The Loeb classical library, LCL 225. England: Harvard University Press.

Ryholt, Kim. 2013. "Libraries in Ancient Egypt." In *Ancient Libraries*, edited by Jason König, Aikaterini Oikonomopoulou, and Greg Woolf, 23–37. Cambridge: Cambridge University Press.

Servies, James Albert. 1950. "Thomas Jefferson and His Bibliographic Classification." Chicago: University of Chicago.

Sickinger, James P. 1999. *Public Records and Archives in Classical Athens*. Chapel Hill: University of North Carolina Press.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

Sider, Sandra. 1990. "Herculaneum's Library in 79 A.D.: The Villa of the Papyri." *Libraries & Culture* 25 (4): 534–42. doi:10.2307/25542290.

Smiraglia, Richard P. 2016. "Foreword." In *Intellectual Activism in Knowledge Organization: A Hermeneutic Study of the Seven Epitomes*, by Hur-Li Lee, xxi-xxxiii. Taiwan: National Taiwan University Press.

Tsien, Tsuen-Hsuin. 1952. "A History of Bibliographic Classification in China." *Library Quarterly: Information, Community, Policy* 22 (4): 307–24.

Witty, Francis J. 1958. "The Pínakes of Callimachus." *The Library Quarterly: Information, Community, Policy* 28 (2): 132–36.

Witty, Francis J. 1973. "The Other Pinakes and Reference Works of Callimachus." *The Library Quarterly: Information, Community, Policy* 43 (3): 237–44.

Zhang, Wenxian. 2004. "Dang An: A Brief History of the Chinese Imperial Archives and Its Administration." *Journal of Archival Organization* 2 (1/2): 17–38.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

Michelle Levy and Christina Orozco

2 Analog to Digital: The Growing Pains of a Religious Archive Migrating its Administrative Collections

Introduction

In the summer of 2013, The Missionary Society of St. Paul the Apostle, also known as the Paulist Fathers, started the process of migrating their archival documents into digital format. The organization decided to begin with a small portion of their records as a trial run; these analog records belonged to a Presidential administration that ended prior to 2013. The first stage of the trial involved the migration of the analog records into digital formats.

The process of migrating an analog administrative collection to the digital world involves trial and error. Attempts were made by the authors to apply web 2.0 principles to the documents but the attempts did not work due to the needs of the community. A number of staff in the office were still accustomed to paper documents. As such, the solution was an adaptation of the organization's specific web 1.0 administrative needs while leaving open pathways for the collection to ease into web 2.0 in the future. Web 1.0 is commonly defined as passive websites (Rubin 2010). The Paulist administrative needs are web 1.0 due to their limited usage of technology during this time, the technology comprising a website and email addresses for different departments. The business documents that are created by the presidential administration are controlled by the Paulist Presidential Board, the developer, rather than the user as would be the case in the web 2.0 environment. At its most basic level, the web 2.0 environment permits site users to interact with and participate in the content development of that site which is an important method for the exchange of ideas (Rubin 2010). Examples of web 2.0 technology include social media sites such as Twitter, Facebook, and Instagram. The Paulist Fathers' ministry, Busted Halo, is an example of the use of web 2.0 technology to interact with users (BustedHalo 2016). The archive however, has not made the transition to web 2.0 yet.

https://doi.org/10.1515/9783110337419-003

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

Background

In an e-mail discussion on November 18, 2016 with canon lawyer Rev. John J. Foley, C.S.P. regarding canon law and best practices in archives, two things became apparent. The first, the Roman Catholic Church provides guidelines for its archives in the Code of Canon Law, Article 2, Can. 482 §1 – Can. 491 §3. Though this section falls under the guidelines for the establishment of diocesan archives, it applies to archives of all religious orders and communities of the church. Secondly, the code is meant to function as a best practice guide for the archives within the church. In *Online Access to American Diocesan Archives: Current State and Lessons for Other Repositories*, Hoelscher's study examines the state of online access and digitization efforts of archives in Roman Catholic dioceses and finds them lagging behind their counterparts. Hoelscher concludes that "all libraries and archives could look at their user groups and their information needs; by identifying these needs and how they can be met, it might be possible to better use limited resources if our target groups have non-traditional information needs" (Hoelscher 2016, under "Conclusion"). Just as canon law written for diocesan archives can be applied to religious communities, aspects of Hoelscher's study can be applied to religious communities. In the case of the Paulist Fathers, the groundwork for digitization is in progress. Addressing the unique needs of the Paulist community while adhering to best practices, both religious and secular, is the goal of this endeavor. Achieving this balance cannot occur without resolving issues that arise from the merger of religious and secular practices.

An issue that quickly presented itself in the early stages of migration of the analog collection to the digital involved the description of the collection. The terminology of web 1.0 and web 2.0 were being used but not the technologies associated with them. This however, is not an unusual approach. Theimer defines Archives 2.0 "[as] an approach to archival practice that promotes openness and flexibility. It argues that archivists must be user centered and embrace opportunities to use technology to share collections, interact with users, and improve internal efficiency" (Theimer 2011, 60). Although the objects would be in digital format when accessed, they needed to be discoverable by the Presidential administration and staff efficiently. Theimer quotes Max J. Evans' *Archives of the People, by the People, for the People* (2007), to clarify that archivists applying Archives 2.0 provide and describe information for their users with tools in a manner in which their users can understand. For the Paulists, this meant that the objects in the digital collection needed to be described using the

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

terminology of their analog counterparts. The result was the creation of terminology based on the colloquial descriptions for the collection.

The use of terminology based on colloquial descriptions, or keywords, was introduced in this project because the Library of Congress Subject Headings (LCSH) did not always have the controlled vocabulary equivalent for the collection. Reliance on keyword searches can cause issues with information retrieval. In Peake's study, which examines the use of Dublin Core (DC) as a minimum metadata standard for Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH), he also examines the use of controlled vocabulary searches versus non-controlled vocabulary searches (Peake 2012). In his testing of the advanced search in OAI-PMH, he found uncontrolled vocabulary searches usually retrieved a larger quantity of results, but the ratio of relevant search results was inconsistent. Searching that used controlled vocabulary according to the appropriate Library of Congress Subject Heading generally resulted in fewer items found by the potential user, but the items retrieved had a 100 percent relevancy rate. While some of the numbers in Peake's (2012) study are inflated due to OAI-PMH drawing from multiple repositories, leading to duplicate results, Peake (2012) noted that "keyword searching is necessary to find all relevant records, but duplicate results and the high number of non-relevant records that are often retrieved do not make this the best search process from a user perspective" (20). Unlike Peake's study, this project involves a religious collection, and one of the primary objectives of this project is to provide access to the older administrative collections of the community by the Presidential Board of the Paulist Fathers. As a result, customized keywords were used in instances where Library of Congress Subject Headings cannot be used.

Digitization of a collection has its advantages and disadvantages as well. As stated by McKay (2003), "digitization in an archival environment includes taking a physical object or analog item, such as an art object, a tape recording, a map, or correspondence" (under "Introduction") and transferring the analog item to a digital medium. Along with other advantages, the creation of a digital surrogate reduces the handling of older or fragile material, extending the life of the original. Digital preservation also faces the issue of storage media deterioration as well as the problem of "rapidly changing storage devices" (McKay 2003, under "Technical Issues"). So that the digital file is not only a surrogate of the analog object, it should also be managed and migrated over time to preserve the digital resource. McKay (2003) also highlights the disadvantages of digitization, which include the cost for "required staff expertise and additional resources... Not only are large budget allocations needed to fund research and intellectual selection, but also time must be spent for feasibility assessments, training, and methodical prioritization of items or collections to be digitized" (under

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

"Disadvantages of Digitization"). This project also faced budgetary and resource issues in order to digitize its collection. Temporary solutions to these issues include the use of affordable technology and software.

Discussion

The Paulist Fathers were founded in July 1858 in New York City by four former Redemptorists priests: Fr. Isaac T. Hecker, Fr. George Deshon, Fr. Augustine Hewitt and Fr. Francis Baker (Paulist Fathers 2016). Since the Paulists are a non-profit religious community, they produce standard business documents such as billing, professional correspondence and personnel hiring. In addition to being a non-profit religious community, the Paulists also behave as a "family" that creates records such as greeting cards, personal correspondence, various insurance papers, funeral documentation, continuing education plans and more. So, the Paulists developed colloquial terms to describe the different aspects of their administrative collection over the 158-year life span of their community.

The community also developed its own method of arrangement for their administrative collection over the years. Each president of the community developed his own method of organization for the papers of the administrative collection, leaving inconsistencies and discrepancies. However, over four decades ago former president Very Reverend Thomas Stransky, C.S.P. created an organizational method that remains in place today. The organizational system stems from the physical organization of the paper documents housed in filing cabinets that contain the administrative collection. The collection is organized into two main categories: church management and administration. These two categories are further divided into four main sections: subject driven files, foundations files, chanceries files, and human resources files which are confidential in nature. Thus, this organizational structure is born out of a web 1.0 world.

In brief, web 1.0 content is the World Wide Web pre-Internet user-driven content; it is static information that is difficult to interact with or contribute to (Aghaei, Nematbakhsh, and Farsani 2012). Web 1.0 mimics analog materials in this way, as it is a static and relatively unchanged method of communicating information. In 2013, the Presidential Board decided that it was time to begin a transition of this system to a digital format. This transition is an ongoing process, with the ultimate goal of making the administrative papers part of web 2.0; via making select records available for researchers online on a case by case basis with approval from the Presidential Board. In contrast to web 1.0, web 2.0 is defined as content found on the World Wide Web that allows users to interact,

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

share and create; the content of web 2.0 is dynamic (Aghaei, Nematbakhsh, and Farsani 2012). The need to work within the confidential nature of the documents in the collection creates a difficulty in finding tools that work within these boundaries.

In order to ensure the human resources aspect of the collection remain private, non-cloud based databases and storage solutions are preferred. Cloud computing relies on the storage of information external to the user's hard drive, often on a third-party server. PC Magazine defines cloud computing as "storing and accessing data and programs over the Internet instead of your computer's hard drive" (Griffith 2016, par. 2). Examples of cloud based storage are Google Drive, Microsoft OneDrive and Office 365. Cloud based content management systems (CMS) include software such as Preservica and CONTENTdm. CMS are defined as "a software suite that is used to create and manage documents and Websites. The term is generally associated with Web content, but systems are developed to maintain all kinds of information" (PC Magazine Encyclopedia 2016, under "Content Management System").

Databases that use cloud storage are not preferred for this project for two reasons: the first is the confidential nature of the human resources documents within the collection, and the second is the lack of a permanent IT staff that might install and maintain a cloud-based content management system for the collection.

Open Source Software (OSS) is defined by the Society of American Archivists (SAA) as "computer code that is developed and refined through public collaboration and distributed without charge but with the requirement that modifications must be distributed at no charge to promote further development" (Society of American Archivists 2016a, under "Open Source Software"). Examples of OSS content management systems (CMS) used for archive digitization projects are Omeka, Greenstone Digital Library Software, dSpace and Collective Access. While there are a wide variety of OSS CMS available, they are often too complex to install and maintain without hiring someone with the necessary expertise. Other CMS available as OSS are easily installed but often times rely on cloud based storage, which conflicts with the confidential aspect of this collection. Furthermore, "out-of-the-box" or ready to install software that is user friendly is frequently not pursued due to their potential cost. The office of the Presidential Board does not have an internal network within which the collection could be stored. This situation has been temporarily remedied through the use of multiple external hard drives and USB drives. While not the most practical solution, the use of multiple external hard drives ensures ease of use and relative security for storage of digitized files from the collection. External hard drives are now available in a wide range of memory capacities and one terabyte to two tera-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

bytes hard drives can be purchased for relatively little expense. Of course, this is in comparison to the cost of an "out-of-the-box" CMS, or the cost of time spent attempting to install an OSS or outsourcing the installation of the OSS to a freelance IT specialist. A set number of external hard drives are used as part of the digitization process and kept within the administrative offices at all times. A set is used for data backup and is regularly updated.

A large portion of the budget for digitization was used for the purchase of an Avision FB6280E Bookedge scanner. While the Epson 1000XL Flatbed scanner is often recommended for digitization, the Avision was selected because its Bookedge feature would facilitate the scanning of documents with bindings during item-level digitization. The current practice is to scan analog materials to Tagged Image File Format (TIFF) and then convert to Portable Document Format (PDF) via Adobe Acrobat to have redundancy in case of file damage. The TIFF is the preservation copy, which the SAA Glossary defines as "the version of a thing set aside to protect its informational contents from harm, injury, decay, or destruction" (Society of American Archivists 2016b). The preservation copy is left unedited as "it is used only to make other copies for access" (Society of American Archivists 2016b, under "Preservation Copy"). The PDF is used as the access file, which is then scanned using Optical Character Recognition (OCR) software through Acrobat. OCR, or Optical Character Recognition, transforms the scanned image into a word searchable document. During the digitization process, the digitized items are indexed and inventoried onto an Excel spreadsheet.

Excel is currently being used as an alternative to cloud based content management systems. Dublin Core is the schema that is used to organize the collection's metadata. Column headers are static and organized by the basic Dublin Core elements. In order to make the cataloging process easier, the data validation function of Excel is being used to introduce drop-down options for the Subject Headings, which is a "controlled vocabulary typically used to provide topical access points for catalog records" (Society of American Archivists 2016c, under "Subject Headings"). This drop-down list of the most frequently used headings speeds up the process of metadata entry since the autofill-suggestion option of Excel does not always work well when inputting large amounts of data. The Dublin Core element dc.abstract is being used to input keywords for the collection. The Subject Headings and Keywords columns are the closest application of web 2.0 for this collection, due to its similarity to tagging. The Subject Headings are constrained by the Library of Congress Subject Heading (LCSH) thesaurus, supplemented by The Oliver Leonard Kapsner, O.S.B. Cataloging Bulletin (Kusmer and Kapsner 2007).

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

In order to fill gaps not covered by either subject heading list, which may occur due to the religious origins of the collection and thus to a very specialized subject list, keywords are devised. These terms are usually colloquial and are used to identify themes that repeat, yet do not have a standard LCSH equivalent. If there is a significant level of repetitiveness the keywords are added to the list of Subject Headings for the collection under the label CFC, or Created for Collection. This is done only if necessary and in each case, they are designed to follow the model used by the LCSH in order to increase the number of relevant search results. Each of these are either approved of or recommended for addition by a member of the community, who has experience in the administrative aspects of the community and canon law, during the quality control stage. "Tags and standardized subjects are not mutually exclusive. [They...] would enable users to follow both standardized and user-tagged subjects; whichever makes most sense to them. In turn, they can add tags to resources. The user responds to the system, the system to the user. This [is...] a customized, user-centered catalog" (Maness 2006, under "Tagging"). Every effort is made to ensure that the organization of the collection meets the needs of the community to facilitate search queries. Current search queries follow the need for known-item search.

The current schema facilitates known-item searches which falls into the classic model of information retrieval. In known-item searches, the user knows either the creator or subject matter of the item. This model is based around the idea of the user presenting a single, static search query from which the user does not deviate, and is completed when the search results match the parameters of the original search query (Bates 1989). This type of search is less open-ended and falls within the web 1.0 framework. The use of tagging through keywords allows the search to evolve or be more exploratory in nature. This means that the initial query evolves to reflect the new information discovered by the user on the path of their search.

While web 2.0 is intended for use on the Internet, its principles can be adapted to a more analog based collection. Combining tagging with controlled classification schemas is one way it has been done so far for this collection. The SAA defines a tag as "a short string of characters used to label the class of a data element in a record" (Society of American Archivists 2016d, under "Tag"). On the web tagging is generally defined by its social media aspects, however it is primarily used to facilitate information retrieval and to give further context to the items being described (Wal 2007). This aspect of tagging remains the same regardless if it is being used on social media (such as a personal or community based blogs, a social bookmarking site, Twitter, etc.), or on a database, or a static spreadsheet. Tagging in the form of keywords creates another way of linking

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

items, expanding the potential for searches that are exploratory in nature. As the keywords used to describe the items are in colloquial terms it is easier to group similar items by topic at a glance, and so it facilitates the retrieval of related topics during the search. While the current schema and system for information retrieval addresses the community's needs, additional solutions are being evaluated to see if they match the needs of the community and offer any advantages.

Conclusion

As the project is ongoing, it is difficult to state with certainty whether the transition from web 1.0 to web 2.0 will be successful. This project is meant to share the issues and solutions the Paulist Fathers archive currently faces while transitioning to the digital world. When keeping the user as the focus, it is necessary to obtain a balance between adhering to religious and secular best practices for the archive. Having the user as the focus of this project leads to the creation of subject headings specifically for the Paulist archive when necessary. It also creates the need to work within a strict set of resources, both financial and technological, in order for the user to retrieve the information they require. However, further studies are recommended in order to better understand the obstacles that unique organizations, such as this Roman Catholic religious community, might face when attempting to shift their archival collections from web 1.0 to 2.0.

References

Aghaei, Sareh, Mohammad Ali Nematbakhsh, and Hadi Khosravi Farsani. 2012. "Evolution of the World Wide Web: From Web 1.0 to Web 4.0," *International Journal of Web & Semantic Technology* 3 (1). http://airccse.org/journal/ijwest/papers/3112ijwest01.pdf.

Bates, Marcia J. 1989. "The Design of Browsing and Berrypicking Techniques for the Online Search Interface." https://pages.gseis.ucla.edu/faculty/bates/berrypicking.html.

BustedHalo. 2016. "About Busted Halo." http://bustedhalo.com/about.

Dublin Core Metadata Initiative. 2017. "DCMI Specifications." http://dublincore.org/specifications/.

Griffith, Eric. 2016. "What Is Cloud Computing?" *PCMag.com*, May 3. https://www.pcmag.com/article2/0,2817,2372163,00.asp.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

Hoelscher, Colleen, 2016. "Online Access to American Diocesan Archives: Current State and Lessons for Other Repositories." Marian Library/IMRI Faculty Presentations. Paper 5. http://ecommons.udayton.edu/imri_faculty_presentations/5.

Kusmer, Robert, and Oliver Leonard Kapsner. 2007. "Oliver Leonard Kapsner, O.S.B. Cataloging Bulletin." *Catholic Library Association*. http://cathla.org/Documents/KapsnerBulletin/kapsnerbulletin1.pdf.

Maness, Jack M. 2006. "Library 2.0 Theory: Web 2.0 and Its Implications for Libraries." *Webology* 3 (2). http://www.webology.org/2006/v3n2/a25.html.

McKay, Sally. 2003. "Digitization in an Archival Environment." *Electronic Journal of Academic and Special Librarianship* 4 (1). http://southernlibrarianship.icaap.org/content/v04n01/Mckay_s01.htm.

Paulist Fathers. 2016. "Our History." http://www.paulist.org/who-we-are/our-history/. PC Magazine Encyclopedia. 2016. "Content Management System." https://www.pcmag.com/encyclopedia/term/40273/content-management-system.

Peake, Michael. 2012. "Open Archives Initiative Protocol for Metadata Harvesting, Dublin Core and Accessibility in the OAIster Repository." *Library Philosophy and Practice (LPP)*. http://digitalcommons.unl.edu/libphilprac/892/.

Rubin, Richard. 2010. *Foundations of Library and Information Science*. New York: Neal-Schuman Publishers.

Society of American Archivists. 2016a. "Open Source Software." http://www2.archivists.org/glossary/terms/o/open-source-software.

Society of American Archivists. 2016b. "Preservation Copy." https://www2.archivists.org/glossary/terms/p/preservation-copy.

Society of American Archivists. 2016c. "Library of Congress Subject Headings." https://www2.archivists.org/glossary/terms/l/library-of-congress-subject-headings.

Society of American Archivists. 2016d. "Tag." https://www2.archivists.org/glossary/terms/t/tag.

Theimer, Kate. 2011. "What Is the Meaning of Archives 2.0?" *The American Archivist* 74 (1): 58–68. doi:10.17723/aarc.74.1.h7tn4m4027407666.

Wal, Thomas Vander. 2007. "Folksonomy." http://www.vanderwal.net/folksonomy.html.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

Christopher Starcher, Joy Perrin, and Shelley Barba

3 The Theory was Sound: A Case Study in the Lifecycle of a Library Streaming Sound Collection

Introduction

In 2005, the Texas Tech University School of Music (SOM) and the University Library reached an agreement to transfer approximately 3,000 classical music Compact Discs (CD) from the SOM Listening Library to the University Library's collection to be used in their listening stations. In a time of budgeting constraints, the move would help the SOM save money while providing the students more available hours to access the recordings for their assignments.

During the same time, the University Library was expanding its mission and focus to provide access to materials beyond the building's work hours in part by developing a digital library infrastructure. Because the collection of musical recordings was specifically a part of the library and because there was a desire to make it more accessible, it seemed like a natural fit to become a digital collection. This chapter will recount the history of the project and then review how common library standards and practices influenced the creation of a collection that ultimately was so difficult for patrons to use it had to be decommissioned.

Planning the collection

A group of librarians came together as the project team to figure out how to handle the complexities of audio files. The team found the complexities of constructing a collection that worked within the legal constraints and the current technology to be an exciting challenge. At that time, only Indiana University had begun to transition away from the traditional listening library model: offering content on physical media, circulated or, most commonly, restricted to only listening stations at a library. The Variations Digital Music Library at Indiana University was a new way for libraries to deliver audio content to users (Dunn and Mayer 1999). However, the Variations' software was then only available in-house, and it would be another four years until it was open for other institutions

https://doi.org/10.1515/9783110337419-004

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

to use in February 2009 (Indiana University 2009). The University Library had to develop its own system and infrastructure for this project, made easier with a $20,000 grant from a local foundation (Starcher and Brewer 2005).

The principal motivation for a shift to this streaming-sound collection project was for the convenience of the patrons (Starcher and Thomale 2008). The idea was to transform a collection that had been historically inconvenient to access into one that would be easily accessible from anywhere in the world for faculty and students associated with the university. This would make the SOM Listening Library available not only to students on campus, but also to those off campus such as distance and study abroad students. To do this well, the user interface had to be simple and intuitive. The project team originally conceived the system to function like a desktop-media application, such as iTunes, which would be familiar to most users. Discovery, metadata display, and streaming functionality would be built into a unified interface, allowing access to all content without the need to switch between interfaces or applications.

The idea of providing students access to these important holdings all hours of the day, seven days a week, was exciting for the librarians. This project provided a chance for them to be innovative and also to push the boundaries of copyright toward fair use of musical content. Fair use is the concept in U.S. copyright that an author or creator has certain rights to their work, but that the general public also has some rights in how they can use that work to create new things. Fair use is a doctrine that protects a limited use of content, such as educational purposes or citing a quote in an article to prove a point. Fair use is not a clear cut doctrine, and ultimately judges weigh if the copyright owners' rights have been infringed. The first issue the team needed to address was how the project would be informed by the legalities of U.S. copyright law. With legal guidance from the University Counsel, the team realized that the system they were designing would have to be able to restrict the use of the collection to those affiliated with Texas Tech University. It was also imperative to prevent downloading and capturing of the bit streams by the authenticated users. Without the necessary system available to meet these specifications, the project team believed the solution was to make a new streaming sound-player that would prevent users from accessing the files directly, which would thereby prevent them from copying the files. If successfully implemented, the file content actually would be more secure from unauthorized duplication and distribution than when it had circulated on compact discs (CDs), which could be duplicated easily on any computer with a compact disc read-only memory (CD-ROM) drive. This requirement meant the project was going to require programming new software.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

When the concepts of the project were presented to them, the library's administration was not as excited by possibilities as the librarians had been. The administration was concerned about the cost effectiveness of a new system when the library had just invested money in the content management system (CMS) CONTENTdm for all digital projects. The two groups reached a compromise: the librarians would use CONTENTdm, which had already been purchased, for the search functionality coupled with extensive metadata creation rather than developing a new search interface. The collection would have a separate streaming-sound player which would stream the audio files securely. The player would only stream the files and display track metadata; there would be no search functionality within the player.

Building the collection

The team identified the need to outsource development of the streaming sound player, as the University Library did not have the resources to create it in-house. Using monies from the grant, the team hired a local development company to develop the player using Adobe Flash. The player communicated with a MySQL database that stored track-level metadata and the location of the files on the file server. The player was designed with the requisite functionality to control an audio stream such as a play/pause button and scrubber bar. The player also displayed the track-level metadata from the database for the current track.

Within CONTENTdm, the search interface for the collection, the links to the player's files could not be a direct path to the location on the storage area network (SAN), a necessity to protect the files from piracy. So, within the system, some additional architecture was developed to prevent downloading of the files. The system processed the request through a series of encrypted data interchanges between the MySQL database and the Flash player, eventually passing the file location securely to the Flash player. This process, controlled by PHP scripts, also included user authentication through the Texas Tech University's single sign-on system, eRaider (Starcher and Thomale 2008).

The SOM Listening Library had been using a DOS-based catalog for many years. Although the records in that system were of high quality, the format of the data could not be easily converted into a usable form for digital files in the collection online nor the physical CDs in the Library's catalog. Although the CDs would no longer need to circulate after the collection became available online, the need for inventory control would remain, and the University Library's integrated library system (ILS) was the optimal system for this task. A survey of

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

the records in the Online Computer Library Center (OCLC) revealed that most of the recordings had MAchine-Readable Catalog (MARC) coded records available. Being able to use the existing OCLC records was highly beneficial. The University Library's music cataloger could add these records to the ILS and then enhance the records with the additionally needed information. This saved time and resources on the University Library's part, and was beneficial on the whole as the newly-rich metadata records were then re-loaded into OCLC for other libraries to use. This process was easy to implement since it followed the workflows and interfaces already established for the cataloging department.

Of the MARC records that existed for most of the collection, many of them were short, and none of them contained the detailed, track-level metadata required for this system. A robust schema was devised to accommodate the needed metadata while conforming to the MARC standard. The Music Cataloger used the MARC 505 field to structure the track-level metadata with standard subfields and punctuation as delineators. This structure enabled programmatic parsing into eXtensible Markup Language (XML) for the greatest compatibility with whichever content management system the University Library would use in the future.

The SOM Listening Library transferred the physical collection of CDs to the University Library, starting the process of accession. After cataloging was completed for all titles, the recordings were housed in locked drawers in the media section of the University Library where patrons could access them with assistance, but not take them from the building. At the same time, the University Library's digitization lab took a portion of the CDs to be ripped in batches. Using Windows Media Player, the lab produced two files for each track. One was a WAVE file to be stored in the archive, and the other was an MP3 file that would serve as the access copy to be streamed.

Initial problems

On paper, putting the digital collection in CONTENTdm worked well. It was a frugal use of money already spent on digital collections. In reality the collection was plagued with problems because the CONTENTdm in the mid-2000s was designed to work best with digitized textual objects, such as books composed of pages, not audio files. In order to give the user the ability to search the collection by track, the metadata librarian created each CD as a "compound object" item. In CONTENTdm a compound object comprises multiple files bound together and displayed as a single item, a function usually associated with a multi-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

paged textual item such as a book. So within the collection, each track was associated as a "page" within that compound object item. This approach was flawed since in classical art music, many works and individual tracks would often have the same name. In addition, the limitations of customization of search results display created situations where a query for "symphony" would produce ten pages of items called "Symphony No. xx" without displaying any other metadata to distinguish one symphony from the other. Searching for the second movement of a particular symphony was impossible. This problem was, ultimately, a combination of how CONTENTdm displayed information and how the information had been mapped from the MARC record; a task that turned out to be more complicated than previously expected (Thomale 2010).

In addition to search functionality difficulties, the project team encountered unexpected problems with collection size. While the University Library had an unlimited license for CONTENTdm, limits within the server file system required the collection to be split. Regardless of the size or type of files put into CONTENTdm, the directory structure would not allow more than 20,000 files in a single directory. Because the digitization of the CDs had divided each track into a separate file, instead of dealing with a few thousand files, the system was having to handle a little over 30,000 files. This forced the metadata librarian to divide the Streaming Sound Collection into two collections thereby making searching even more confusing to patrons. In order to search both online collections, patrons had to search all of the digital collections, resulting in even more disambiguation problems.

Unknown to the project team, these early problems would not be easy to overcome on the users' side. While the idea of the collection at first had excited the music faculty, they soon lost interest once they experienced how difficult it was to find the titles they would assign to their students. The collection languished, was poorly used by faculty, and was mostly ignored by the University Library once most of the original project team had moved beyond the project or left the library completely.

Assessment and reworking the collection

With a bit of restructuring and reorganization at the University Library in 2012, including the formation of a Digital Resources Unit, fresh eyes turned to all the digital collections to evaluate what had been successful and what had not been over the previous seven years.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

The Streaming Sound Collection was an obvious choice for reworking as time had only made the interface seem clunkier. The new group conducted a user experience audit, surveying, interviewing, and observing the music faculty using, or trying to use, the collection. The music faculty confessed to actively telling new faculty and students not to use the Streaming Sound Collection from the very beginning because it was so difficult to navigate.

There were a few different ways the collection could be fixed. First, something could be done with the titles to display more metadata associated with each track by combining MARC record subfields. However, reworking the metadata would require processing the MARC records again, and the new team did not have the skill set nor the resources to do so. Instead, the Digital Resources Unit devised a simpler option: the collection could be reformatted so that the CD-level metadata would be made into an "item" with the track information included in it. As a proof of concept, the digital resources librarian set up a Mediawiki instance and reformatted the metadata from CONTENTdm to display the CD-level metadata with the streaming player tracks linked on the page. Then, she put a small number of titles into the system and used linked data concepts to enrich the displayed metadata record. The unit thought the extra benefit of being able to link any part of the page, including composers' names and instrumentation, to other pages would allow users the ability to search and browse easily.

Initial user tests with students of this new collection showed a marked improvement in usability and search results. The Digital Resources Unit then reintroduced the collection in February 2012 for a testing period with music faculty and students to determine if the initial problems were fixed with this reorganization. During this trial, the original collection remained linked on the University Library's website, while only the music faculty were given access to the smaller test collection's location. Using Google Analytics as well as the web server logs for the tracks in the Streaming Sound Player, the digital resources librarian monitored usage statistics on the collections in CONTENTdm and Mediawiki. The trial period was from February 29, 2012 to April 23, 2012.

During the trial, the collection saw a resurgence of usage and interest in the Mediawiki instance. Faculty and students viewed the collection more often, and even viewed a wider breadth of the collection. Whereas before in CONTENTdm only four percent of the collection was used, during the trial test 72.93 percent of the collection was viewed. To analyze the number of titles that users actually were listening to, the digital resources librarian pulled data from the web-server logs for the streaming sound player. A hit on the streaming sound server meant that someone had clicked on a link, logged in with their credentials, and listened to at least some of the track.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

During the trial period, the files were accessed a total of 6,192 times with an average of 109 views per day. The average view per visitor was 5.43, meaning people were listening to five to six tracks at a time. The server had a total of 1,112 visitors with an average of twenty visitors per day. In 2011, during the same time period when the collection was in CONTENTdm, users had listened to only 657 tracks.

This data seemed like a vast improvement and an indication that the new interface had fixed the previous usability problems with the collection. The group finished moving the whole Streaming Sound Collection to the Mediawiki and linked it from the University Library's web page. Not only did this seem to be a better content management system for this particular collection, but the library's administration had decided to move away from CONTENTdm, and it was retired in 2013.

In mid-2014, the Information Technology (IT) team in the University Library alerted the Digital Resources Unit that there was a problem with the Uniform Resource Locators (URLs) used for the Streaming Sound Collection in Mediawiki. The Unit had not been aware that the linked URLs in the collection contained embedded references to the CONTENTdm server. Since CONTENTdm was no longer being supported, the links had been broken for a few months. A red flag that something was off with the collection was that this information came from the University Library's IT team, not the music faculty. Since it was early in the semester, there was still the chance that the music faculty had not yet assigned the tracks for the students to listen. Additionally, the music faculty also had a history of not being vocal in their complaints about services. So, the Digital Resources Unit assigned a few student assistants to manually correct the URLs, and continued on. While this fixed most of the broken URLs, another audit at the end of the semester revealed that thirty-six percent of URLs were still referencing the CONTENTdm server. This was troubling to the unit because no one had reported that the collection had not been working for the whole semester. Hearing that more than a third of a collection had been unusable for more than a year and no one seemed to mind caused the unit to re-evaluate the worth of keeping the collection.

Closing the collection

In part, because of the absence of alarm when the collection was broken, and partly because of an off-hand comment from a music faculty member that it was easier to find listening assignments through free online services such as

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

YouTube or Spotify, the metadata librarian who had been working on the collection decided to conduct an assessment between the semesters to see how many of the items in the Streaming Sound Collection were available through other services. A student assistant was provided with (1) a spreadsheet, (2) a list of CDs and tracks contained in the Streaming Sound Collection, and (3) instructions on how to search both Spotify and the library's subscription music-databases such as Classical Music Library. In checking a sample of 3,293 tracks from the collection, 1,190 (thirty-six percent) had exact matches in Spotify, and 2,015 (sixty-one percent) had matches that were the same work in either Spotify or Classical Music Library, while not the exact performance. In fact, only thirty-four (one percent) tracks could not be found in Spotify or in any of the library's subscription music-databases.

After consulting the music liaison librarian, the Digital Resources Unit could confirm that only two music faculty actually were using the collection for their classes, and only one was assigning a few CDs out of the over 3,000 available. The original justification for digitizing this collection under fair use, that the library was providing better access to difficult-to-obtain educational materials, was questioned since the vast majority of the collection was not being used for classes and the content was easily available through subscription services. How could the library justify holding copies of thousands of CDs if (1) they were not being used for classes, and (2) the content now was available through other means? The library no longer had a strong claim for keeping a copy of the copyrighted files since most of the files were not being used consistently by classes and many of the tracks were available other places.

To put a nail in the coffin of the collection, problems with the player itself became insurmountable. In late 2014, the Streaming Sound Player broke after an update to Flash. While the Library's IT team was able to get the player back up and running after much difficulty, the process highlighted how difficult the technology was to maintain. If the collection were going to continue, there would have to be an investment to update the player. A few months later, the Flash vulnerability became big news (Adobe 2015). With all the discussion in the technology literature about the impending doom of Flash, it seemed the Streaming Sound Player itself would be outdated in just a few years. The library could no longer guarantee that the custom-made Flash player was going to work, and the chances of finding the funding to upgrade the player were slim.

Everything was falling into place to deaccession the collection, and on September 1, 2015, it was no longer available for users to access through the Library's website. When sun-setting a service like this, even though the usage was not great, the library made an effort to alert the two faculty members that

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

had been using it, and kept the handful of tracks available that were directly linked from their course materials that semester.

By the end of the year, the Digital Resources Unit had deleted all copies of the CDs from the server. Nearly ten years after the project began, the Streaming Sound Collection had died. While the collection had some problems from its inception, it was the execution of the project that fed those problems, creating an informational resource that had become too costly to fix compared to its usefulness. A digital collection that is difficult for patrons to actually use can only be a source of frustration for all involved.

Lessons learned from the collection

No matter how well researched and planned a digital project is, it will always fail if the actual product is difficult for patrons to use. In the case of this project, all of the components of the project plan, from creating the MARC records, to ripping the CDs, to converting the MARC records to metadata records for the items, and then creation of the Streaming Sound Player, worked smoothly. Every part of the project was efficient and careful, and yet the end product did not work because of an incompatibility between the information needs of the users and the system housing the items that they wished to retrieve. The earliest mistake in this project was the choice to force the collection to fit into the structure of CONTENTdm rather than its own system. This choice, made by upper level administration, did not work with the rest of the project and as implemented made the collection mostly unusable for most of its life. The correction in 2012 was too little, too late, as the whole ecosystem of streaming music had changed. Sometimes, the system choice really does matter, and an organization cannot shoehorn all projects into a single content management system and expect them all to work to their best potential. Sometimes, special collections really do need their own organization and presentation platforms.

The first lesson learned was to avoid the *sunk cost fallacy*. The sunk cost fallacy is when an investment of either time or money makes those involved with a project want to continue just so that their previous investment will not got to waste. This fallacy is particularly dangerous for expensive technology projects that cost money to continue. The library had spent a lot of time and energy on the Streaming Sound Collection over the years, but in 2015 it was no longer worth the effort to keep it going. A few people did not want to shut off a service, nor give up on a collection that had taken so much work, but keeping it going would have cost too much with little in return. All digital collections have

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

a lifecycle, and ones that depend heavily on multiple specialized programming or a specific format (e.g. Flash) have a shorter cycle than most. A once innovative service can quickly become outdated as slicker commercial or open-access options become available. In the case of this project, even in the best scenario, the Streaming Sound Collection was only going to be useful until a better streaming service or database was made available. For this project, the creation and success of Spotify, launched in the United States in 2011, was one of the major reasons why there was no longer a need for the Streaming Sound Collection. Even if everything had gone perfectly and the Streaming Sound Collection was heavily used, there would have still been a disruption to the project when Spotify became available, and the viability of the project would have been called into question. The collection, in this case, potentially only had a lifespan of one to two years before it was to become obsolete.

The second lesson learned was to identify failures more quickly by developing innovative projects in phases. Early testing possibly would have prevented the usability problems this project experienced. If a small batch of items had been done, and presented to the faculty, the feedback would have been easier to gather and act upon. Instead of faculty telling each other not to use a service later on, the project team could have taken steps to make the collection more usable. While the project team did test the project, only librarians and IT staff beta tested and gave feedback, not those who would eventually (try to) use the collection. Along these lines, scope should be carefully guarded and strictly controlled with technology-heavy projects. Here, the whole collection would have been more manageable and more fixable if instead of scanning the whole SOM Listening Library of 3,000 CDs, the project team had only scanned the items that were to be used during the upcoming semester. This would have produced a useable product more quickly, and because of the smaller scale, the problems would have been more easily identified and fixed. It also would mean that the process could be refined for the rest of the CDs. Given that the collection had a limited potential life, this would have also make the collection available earlier. A focused scope would have also helped to ensure that the items being ripped were actively being used. This would have strengthened the fair use case for having the items streaming in the first place.

The third lesson was to build assessment into technology heavy projects and make a separate unit responsible for it. If the assessment of how many tracks were actually available through other means had been done earlier, the project may have ended sooner, and it would have saved time and resources. Often, project teams are so often burned out from completing a project that assessment becomes an afterthought. Assessment requires dedicated energy and time to be done correctly. Second, for a project team it can be demoralizing once

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

you finish a project to critically see that it is not working. Those heavily involved with the project's implementation may have problems developing the perspective to rethink how things could be done better: they may no longer be objective. It is important to balance the assessment and maintenance of old collections with the creation of new collections. While discussions were happening about how to fix the initial problems, a number of solutions were raised, but implementing them would have prevented work on other newer collections. The time between the project's launch and the eventual re-working was too long, but in the moment there was always something more pressing and the collection was often pushed off to the side. When projects are heavily dependent on specialized technology, if the service is to be maintained, then system maintenance needs to be given priority above other new projects. With the hindsight that the collection only had the potential of a two-year life online before it was likely to be obsolete, it is easier to see that a quick fix to the track titles (no matter how crude) would have been better because the collection would have been at least more useable.

Conclusion

As the University Library designed other digital collections, the direction of the Streaming Sound Project became influenced less by the vision of the team and more by organizational decisions outside the scope of the project. These decisions led to the development of an infrastructure that placed institutional priorities ahead of user needs. The collection was placed in a system that was unsuitable for the content and robust metadata, and the project was concluded.

The resulting collection was difficult to navigate and usage never picked up. After five years, a new team tried reformatting the collection's metadata and migrating everything to a new system. This system, while much more suitable and user-friendly, debuted just as independent streaming services such as Classical Music Library and Naxos Music Library began offering well-designed user interfaces, robust catalogs that surpassed the content of our local collection, and competitive subscription prices.

While the project was not a complete failure, it is an example of library standards and practices influencing the creation of collections and systems that are difficult to use. The fear of failure, combined with competing priorities, limited resources, and sunk-cost thinking, stifled the collection. Through the telling of this case study, the authors hope to help others to better plan and execute innovative projects and metadata practices at their institutions while also

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

realizing that like other library holdings, the digital collections are only temporary.

Digital collections, especially innovative ones, are not permanent. If those involved had gone in expecting that the collection would only be useful for five years, they would have changed how they handled the problems that came up or found a way to fix problems faster.

References

Adobe. 2015. *Security Advisory for Adobe Flash Player*. https://helpx.adobe.com/security/products/flash-player/apsa15–04.html.

Dunn, Jon W., and Constance A. Mayer. 1999. "VARIATIONS: A digital music library system at Indiana University." In *DL '99 Proceedings of the fourth ACM Conference on Digital Libraries*. August: 12–19. doi:10.1145/313238.313242.

Indiana University. 2009. "Indiana University Offers Open Source Digital Music Library Software." *Indiana University Libraries Digital Projects & Services*. February 18. http://www.dlib.indiana.edu/projects/variations3/docs/variations-oss.pdf.

Starcher, Christopher, and James C. Brewer. 2005. *Streaming Audio Grant to the Helen Jones Foundation*. Grant Request, Lubbock: Texas Tech University. http://hdl.handle.net/2346/508.

Starcher, Christopher, and Jason Thomale. 2008. "To Stream the Impossible Stream." *Music Library Association 77th Annual Conference*. Texas Digital Library. http://hdl.handle.net/2346/506.

Thomale, Jason. 2010. "Interpreting MARC: Where's the Bibliographic Data?" *Code{4}lib Journal* (11). http://journal.code4lib.org/articles/3832.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:06 AM

Dara Lohnes-Davies

4 Digital Access Enhancement Initiative at the National Music Museum

Introduction

This chapter will describe the strategies and processes that the National Music Museum (NMM) in Vermillion, South Dakota implemented in a high volume digital enhancement initiative. The two-year project involved photographing and post-processing 3,600 musical instruments from the museum's founding collection. Because the NMM has not attempted a project of this magnitude before, especially under the give time constraint, this initiative burdens the small museum's resources. With projects of this scope, generating easily accessible and intuitive information while maintaining efficiency requires foresight and attention to detail. Documenting these strategies and processes, along with successes, challenges and lessons learned is crucial in order to understand what is needed to make the entirety of the collection available to the global audience while maintaining best practices.

Taking on a large project in a small museum is like trying to construct monumental architecture. Three examples come to mind; Saint Petronius's Basilica in Bologna, Saint Mark's Basilica in Venice, and Saint Peter's Basilica in Rome. These three buildings have some similarities. They are famous works of Italian architecture, are all places of worship, are popular tourist destinations, and are each an ambitious project in their own right. They differ in historical context, implementation of plans, and final results. The modern Saint Peter's construction was started in 1506 and completed in 1626. It is idolized for its large size, beauty, and mastery and a place that people around the world dream of traveling to experience. It is like a large museum, working on a project with many resources and people to help it come together. Compared to Saint Peter's construction, Saint Mark's and Saint Petronius's represent two methods of approaching a similar sized project for smaller museums. Saint Mark's construction started in 828 and was completed in various stages, continuing the project through to the 1500s. Saint Petronius's was started in 1390 and through various events is now more famous for never being completed despite its high aspirations. Essentially, too much was added in one project to the detriment of its completion. These same concepts are relevant to the museum field. High ambitions paired with not acknowledging limitations can result in damage to collec-

https://doi.org/10.1515/9783110337419-005

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

tion objects, declining audiences, or worse. The author's experience in undertaking a digital enhancement initiative at the National Music Museum has proved the following: that careful planning and attention to detail is necessary for any ambitious museum project to be a success. The following chapter documents the processes the staff of the National Music Museum developed as a part of the aforementioned digital enhancement initiative, and the effects of these processes on the museum's approach to ambitious projects. By documenting the project in this manner, the hope is to help museums of similar size and staffing avoid reinventing the wheel when planning projects, so that others can successfully meet their own goals.

Project overview

In February 2016 the National Music Museum (NMM) in Vermillion, South Dakota began a digital access enhancement initiative for two subsets of its collection. These subsets are the founding collection of Arne B. Larson and the collection of pre-1800 instruments, totaling approximately 3,670 objects altogether. The two-year project involves updating the catalog, creating a two-dimensional (2D) image package, and making the resulting information available online with a website overhaul. This chapter focuses on digitization efforts and photography workflow, documenting the strategies and processes the staff has implemented. As the NMM has not attempted a project of this magnitude with the given time constraints, the staff has encountered some growing pains. Subsequently, they have reevaluated the institution's precedents to make the project more manageable for the current staff and resources. With projects of this scope, efficiently generating easily accessible and intuitive information requires foresight and attention to detail. It is crucial to document this effort's successes, challenges and solutions in order to understand how to make the entirety of the collection available to the global audience while maintaining best practices as a small museum.

As this project was still in progress at the time of writing, it will only provide preliminary results. The author first describes the NMM's collections and the project's scope. The photography workflow rationale follows, with a short analysis of the sources consulted. A description follows of the workflows that had been tried, as well as a discussion of time costs and results from those initial workflow trials. The conclusion discusses the surprising benefits of this project and the project vision for its future. By presenting all of this as a case study, the staff can help similarly sized institutions plan their own digital ac-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

cess enhancement initiatives. Although this chapter is not a "how-to" guide, it provides ample information on the processes that can make similar projects successful.

History of the National Music Museum and its collection

The National Music Museum was founded in 1973 at the University of South Dakota in Vermillion, with instruments from the collection of Arne B. Larson. Larson was the father of the first museum director, Andre Larson. Initially, the collection consisted of assorted band and orchestral instruments. Because both Arne and Andre wished to create an encyclopedic collection, all instruments were given a home in the museum regardless of their origins and physical state. Today the collection has grown to over 15,000 musical objects from all regions and time periods. It is renowned for having one of the finest and most comprehensive instrument collections.

Initially, the NMM had no collection policy so all instruments were accepted. The founding collection, while eclectic and charming, now presents multiple conservation and storage conundrums. Many objects are not stable enough to be displayed or studied, and in some cases, objects are combinations of parts from other objects to create a new whole. The museum's current collections policy helps staff avoid these issues today, but the instruments in the initial founding collection are still a significant part of the collection and require attention. While the organization would not accept much of Arne B. Larson's collection currently, it is a manifestation of the passion and personality of our foundation. Few of these objects, about thirteen percent, have been photographed, and the photos themselves are mostly black and white negatives. This project aims to digitally document this core collection of 3,357 objects plus 352 pre-1800 objects and make them available online.

Determining the full photographic scope of the project

Before photographing, staff had to determine the full scope of the project. The team wanted to plan a workflow based on object type and size, so the list of

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

objects in the proposal was placed in a spreadsheet, then sorted them into tabs which included tracking formulas so staff could track progress. Then each tab was sorted alphabetically according to instrument name, to group like instruments together. The project team wanted to further sort the instruments by construction materials, such as metal, wood, and plastic for clarinets, but catalog inconsistencies prevented them from doing so. Sorting by storage location was found to be the best approach, as it would allow for them to plan the order of the project by instrument type and location.

After the initial sorting, it was important to determine the quality of the extant photography. Since the plan was to publish the images online with basic catalog information, the photographs need to be standardized. The photographer generated a list of requirements that each legacy photograph needed in order to be used, while allowing staff to use a majority of the existing photographs. While the questions could be easily answered they did not address the quality of the images such as sharpness, resolution, and file extension. These questions are important because the legacy photographs have varying degrees of sharpness, differing resolutions and file extensions. For certain objects, we elected to delay photography until the end of the project. The final project list was generated into a titled Master List. Columns were added to track when an object has been cataloged, when the photograph has been loaded, and when it has been approved for publication. It goes without saying that the scope of this project is varied, vast, and challenging.

Deciding on a workflow

From the start the team knew that the proposed workflow in the grant would not work with the current personnel. The grant was written with a different group of people in mind, and by the time it was implemented those people were no longer part of the project. According to the grant, the Associate Curator was tasked with taking high-resolution documentary photographs of each instrument while the Photographer took publication-quality photographs of the pre-1800 instruments. Because the new personnel had different skills and expertise, the grant workflow was altered accordingly.

According to the master list, this project has many small categories of similar objects. Staff wanted to establish a workflow that tackled similar objects, but was varied enough to maintain the team's mental efficiency. To create a workflow process to propose to the project team, the photographer consulted a variety of industry standards. The five standards found to be the most helpful

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

were the (1) AIC Guide to Digital Photography and Conservation Documentation, (2) Musical Instrument Museums Online Photography Guidelines (MIMO), (3) Digital Program Planning (DPP) by Digital Transitions, (4) the Federal Agencies Digitization Guidelines Initiative (FADGI), and (5) the Universal Photographic Digital Imaging Guidelines (UPDIG). Concepts were incorporated from each in order to create the project's digital enhancement workflow.

AIC Digital Photography and Conservation Documentation

This document provides a basic workflow structure that can be adapted to any project in its first section. Staff was able to use this as a starting point for forming the workflow (Warda 2011).

Musical Instrument Museums Online Photography Guidelines (MIMO)

This guideline is specific to musical instrument photography, serving as a guide for creating object sets based on instrument type. However, these guidelines provides little practical information for safely setting-up instruments for photography. Section thirty-six of the document suggests that each photographic session has a precise list of objects to be photographed. The photographer established a 'weekly photo tracking' document to use for this purpose (Bar 2011).

Digital Program Planning (DPP)

The DPP examines the mental factor of large digitization processes. It proposes implementing structured break times for teams and creating time blocks when distractions are kept to a minimum (Digital Transitions Division of Cultural Heritage 2015). Section three is useful because it focuses on pre-planning workflows. It elaborately discusses all the parts of the storing digital assets, how to manage provenance and copyright, and how to prioritize the parts of the collection to be digitized first (Digital Transitions Division of Cultural Heritage 2015). Section four contains workflow suggestions focusing mostly on capturing raw

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

images and processing them efficiently (Digital Transitions Division of Cultural Heritage 2015). What made the DPP document the most useful for the NMM's project is that it combines theory, practical instructions, and case studies that show average working times based on workflow style.

Federal Agencies Digitization Guidelines Initiative (FADGI)

FADGI discusses the ideal workspace for photographing en masse, which was helpful for thinking about setting up off-site storage spaces (Federal Agencies Digitizing Initiative – Still Image Working Group 2015). Another helpful section contained two workflow plans for large scale projects (Federal Agencies Digitizing Initiative – Still Image Working Group 2015). Workflow one was most suitable for this project as it encompasses projects with an established collection and end date, although the photographer also used parts of workflow two because it covered mixed object workflows. The theory discussed in the FADGI that was most helpful when presenting the proposed workflow to staff was the perspective of Adjusting Image Files (Federal Agencies Digitizing Initiative – Still Image Working Group 2015, 60). FADGI holds that cultural heritage files need image processing to produce good, publishable images. FADGI also discuss storing untouched raw files along with edited deritiviatives and explains the ways in which this needed storage is cost prohibative to smaller institutions.

Universal Photographic Digital Imaging Guidelines (UPDIG)

The UPDIG document categorizes four different workflow types and analyzes requisite end products and object types. UPDIG categories consist of (1) high volume, quick turnaround; (2) high volume, moderately quick turnaround; (3) low volume, high quality; and (4) high volume, high quality (UPDIG Coalition 2008). This rationale influenced the approach taken to this project, particularly in reconciling the high volume, high quality work desired with the high volume, and quick turnaround needed for success.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

The team began the project with a series of tests which are described in the workflow trials section of this chapter, with various instruments and localities to gather data. It was decided that after six months the team would stop to evaluate the processes, and determine which method would yield the closest results to the project ideal and still be the most efficient. This evaluation was in ongoing at the time of writing.

The workflow trials

The big concept for the project workflow was to establish multiple curatorial workflows feeding into the photography lab in a staggered and simultaneous method. The curator would act as the funnel for the objects being brought to the photographer from graduate students working as interns on the project. The reasoning behind this workflow was that the file names were dependent on the students' cataloging for information. The students would submit weekly lists of objects completed for the curator to review. Those objects were compiled into a list, guiding the workflow for the next week. The lists also served as documentation to track the project's progress. As the post-processing was completed, the images were then passed along to the next stage to be loaded into the database, and then noted in the Master List that the photograph was ready for publication.

There were three different locations that the team wanted to use to experiment with the photography workflows. The team decided to choose objects with which they were familiar as the test object packages for these three localities. This would allow the majority of the team's focus to be on the workflow between the individuals working on the project and not the objects themselves. The first object package was brought to the photo studio to be photographed. Most of these objects were from off-site storage, so staff spent a great deal of time transporting the objects between off-site storage and the photo studio. The studio is the only location where the team could utilize the tethering technique, due to not having mobile equipment such as a laptop capable of being moved, and as a result could directly sort files and name them in the imaging process. Utilizing this technique saved a significant amount of post-processing time and reduced mistaken identities when sorting the files.

The next location was in the institution's largest off-site storage facility. A hurdle for this set-up was the limited amount of space. Due in part to those restrictions there was a slight drop in the quality of the images being produced. As the NMM currently does not have the capability to accommodate off-site teth-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

ered photography, staff had to rely on manual lists to track which objects were photographed. This increased a margin of error during post-processing filing when images were transferred from the camera to the computer. However, one advantage of this set-up was that it drastically reduced the amount of time it took to move the objects to be photographed.

While the team was photographing in our larger off-site storage, a student and the curator worked on an object package in the smaller off-site storage space to test the stacked workflow approach. This smaller location required the photography studio to be set up and taken down for each session. This added approximately forty-five minutes of work time to accommodate the workflow. Similar to the larger off-site storage shoot, time was gained by not having to transport the objects, but lost in the set-up accommodations as well as in the post-processing because of the need to sort and rename the files after taking the larger batch.

Overall, the workflows appeared to balance each other in terms of time expenditures except for in the latter scenario. The team then had to decide where to spend the bulk of their time; either in transporting objects from off-site to the photo studio, or in taking the studio to the objects, which sacrificed some image quality. The NMM is now considering the purchase of a laptop and affiliated software to make the tethering process mobile, which would eliminate the file sorting aspect of the post-processing. Until then, they have decided that it would be best to transport the objects to the photo studio for the imaging unless the objects themselves make the process difficult.

Time costs: Prioritizing and budgeting time

Other tasks were added to the project, which were not a part of the initial proposal. These included, on the photography side, taking detailed shots of each object beyond the proposed documentary shot as outlined by the object proposal and updating the packaging of objects before they would be returned to storage. Neither task is difficult, but both require time to accomplish according to best practices. The team has struggled to balance these important additional tasks with the previously outlined project proposal. This issue has no straightforward solution and the team continues to strive toward balance in this regard.

Much of the discussion during the project team meetings centered on the question of depth versus breadth. Specifically, will the project be completed by going into complete detail on as many objects as is possible within the time of the grant disregarding the amount of objects proposed, or should the time re-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

sources be focused on the attempt to look at every object at a more simplistic level. The answer to this question revolves around (1) what the team proposed it would do as a guideline, (2) what the team ideally wanted to achieve with the project, and (3) what was manageable in the time allotted to complete the project. Overall, the team has to determine what the NMM's definition of 'complete' is for this project. From the photography standpoint, this is a digitization enhancement project. That means that if the team does not successfully capture at least one image of each object, it has not accomplished what it has set out to do. In contrast, the institutional precedent is to take detailed photographs of an object in one-off projects. More broadly, the team has struggled with balancing ideal results against the project's constraints. The following describes each side of this issue and the steps the team has taken to balance project idealism with practical considerations.

Depth wins on the intellectual and best practices side. Every object should ideally be given equal photographic and conservational treatment. Unfortunately, we do not have the staff time to adequately manage the high volume of objects in the project. It also creates an environment where individual objects within the collection do not get looked at on a regular basis due to spending a lot of time delving deep into each object once. Shifting our focus to breadth is healthier for the collection as a whole, as objects will get attention more often and issues with them can be addressed quickly. The best conceptualization of this is to think of managing the collection in concentric circles; each project you take on is one circle. By focusing on breadth, not only are the objects observed more frequently, creating more opportunity to catch conservation issues, they are all made accessible to the public at least at a basic level. The heart of the NMM's project is to make an entire collection accessible for online consumption. Our focus on in-depth examination of individual objects is holding us back from fully realizing this goal. While in-depth examination will always be important to the NMM, we will also begin experimenting with ways to cover the breadth of our collection, making it more accessible and useful.

What we have learned from our initial workflows

The lessons we have learned from this project are that (1) it is crucial to define the entire scope of any project before beginning it, (2) ambitions must be tempered by practicality, and (3) more specific and smaller goals breed better results on the larger scale. Though we began with the broad idea of having all of the founding and pre-1800 collection imaged and cataloging updated for publi-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

cation, we did not have a complete grasp of the specificity of what it would take in order to organize and make it feasible in our two-year timeline. We knew we wanted to completely digitize this collection, but we are still wrestling with what we define as "complete" as we work through the project. While we cannot anticipate every issue that may come up throughout a project, having a specific, detailed plan helps us to deal with contingencies as they arise.

Another lesson we learned was that tasks within the project should remain independent when possible. For example, we chose object sets for our initial photographic tests based on the amount of cataloging that had been completed with them. However, these object sets required substantial adjustment of the photographic equipment, slowing our photography. Photographing objects based on size and type would speed the photographic process, but hinder the cataloging process. Our file-naming convention further complicated the issue, as it is based on catalog data. In the future, it will be even more important to not only do similar objects as overall sets, but to choose flows that lead to each other more directly to speed the photographing process. By focusing on broad categories from the Master List rather than subsets, equipment adjustments would be less drastic, making the whole process more efficient. We have also determined that having a file naming convention that is directly tied to our cataloging prohibits the photography workflow from moving at its own pace. In the next quarter of our project, our file naming convention will no longer depend on having catalog data completed.

For the sake of efficiency, we are currently looking at reducing the amount of images we take from a complete object package with quantity determined by the type of instrument ranging six to ten photographs to an average of three for each object. We at the NMM knew that the project was large and that we were setting high standards for ourselves to achieve beyond what we initially proposed. Armed with metrics gathered from our initial tests which demonstrate how much time it would take to construct complete object photographic packages in the various localities, we can now show the practicality and feasibility of our future project proposals. Reducing the amount of images taken will make our project more feasible. For example, we plan to eliminate all detail images from this initial workflow because it requires more time and equipment adjustments than our timetable will allow. If the camera tripod has to be adjusted between shots, that is a small amount of time added to our workflow. Trying to do this fast is detrimental to equipment and quality of the end products. Also, reducing the amount of shots for each object will greatly help. This will facilitate this project's focus on breadth, exemplifying the circular approach to photographing the objects. Objects would be handled more often to get more photography as needed as a result, and our photographic emphasis would then closely

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

match the demands of our audience. This level of photography also informs those decisions more readily in-house and when they are released to the public as a whole.

Another aspect of consideration is to have more specific and practical goals in order to breed better results overall. A project of this size should ideally be broken down into more manageable segments that add up to complete the overall project. This would boost the morale of those working on the project. It also helps create data that is useful for making further proposals on similar projects. Ultimately, we have learned that ambitious projects like this require a great deal of planning, and that we must temper our idealism with pragmatism. We began with a great deal of ambition and a basic plan, but found that a more planned and practical approach was necessary. As a result, much of the first six months of our project have been experimental, trying many approaches to set different precedents and establish enough data points to readjust our proposal.

What have we done well and the benefits that we have received so far?

The most useful result from this project is that it forced us to examine the institutional precedents that have been set and analyze how practical they are with our current staffing and goals. One of our biggest setbacks is related to our digital preservation capabilities. We are creating more objects in our digital archives and it is forcing us to think about how we are storing them and making them accessible. Until we are able to determine a more stable solution these images are accessible on the universities server. We have come to realize that these digital objects are not only assets or tools for our collection, but objects themselves. Thus we need to give them the same preservation rights as their physical counterparts.

This project has also forced us to analyze our physical storage and accessibility to our collections. We have a vast collection of objects, and like most museums, only a small percentage is accessible to the public. This project has shown us just how much work will be needed to bring our collections to the absolute best up to date best practices.

We have also found that we can give students a wonderful opportunity to work with our collections in a hands-on way, which directly contributes to our work around the museum. Because the NMM is on a university campus, and is also a world-renowned center for studying instruments, we have a wide array of

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

students with various levels of expertise in organology and museum studies coming through our doors. We have been able to give these students substantial experience of best practices in handling objects, cataloging, digitizing, and post processing our digital assets. This digital enhancement initiative has thus greatly enhanced their education, in a way that may not have been possible otherwise.

The biggest asset that this project has given the photography studio directly is the opportunity to create in-house resources and guidelines for digitizing incoming objects. Our founding collection covers at least briefly, the bulk of different types of objects in our collection. We have been tracking and notating various metrics, such as the types of images are taken of a given object at various levels, the number of images taken, the general studio set ups for each object type, the resources that were used to photograph them, and the average time it took to photograph that specific type of instrument. Post processing has not been notated as that changes quickly and is generally identical from object to object. This is an attempt to categorize and give time estimates for photographing the various objects in our collection. It will ideally provide a foundation for seeing how long future projects will take so we can plan our timelines accordingly.

Where we hope to go from this point in our project

While much work remains to be done in this project, it is important to keep in mind that this project is still in its initial stages. Now we have strong specific data telling us just how much and what is feasible for us to accomplish. Much of what we have done is gather metrics and establish what can be done in a given time frame. With this information, we can aim for a more efficient and successful workflow, which can be managed within the constraints we have given ourselves. We have established guidelines for our own institution, as how to manage high digitization workflows and the staffing needed to eventually get the collection completely digitized. By gathering this data, we will be well prepared for our next digital enhancement initiative. Ultimately, our project serves as a case study for other small to medium museums to examine when planning their own digital enhancement initiatives. By incorporating our ideas and observing our trials and errors, other small museums can be more like a successful

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

Saint Marks as compared to a Saint Petronius, when they aspire to complete Saint Peter sized projects.

References

Bar, Frank P. 2011. "The MIMO Digitisation Standard: Defintion of Scanning Properties and Recommendations for Photographing Musical Instruments." *The MIMO Digitisation Standard*. Musical instrument museums online, August 29. http://www.mimo-international.com/documents/MIMO_Digitisation_Standard_v3.pdf.

Digital Transitions Division of Cultural Heritage. 2015. "Imaging for the Future Digitization Program Planning A Comprehensive and Practical Overview of Cultural Heritage Digitazation." *Digital Transitions*, August 10. http://dtdch.com/ImagingFortheFuture-DigitizationProgramPlanning.pdf.

Federal Agencies Digitizing Initiative – Still Image Working Group. 2015. "Techinical Guidelines for Digitizing Cultural Heritage Materials: Creation of Raster Image Master Files." *FADGI Digitization Guidelines*. Edited by Thomas Rieger, September 2. http://www.digitizationguidelines.gov/guidelines/FADGI_Still_Image_Tech_Guidelines_2015–09–02_v4.pdf.

UPDIG Coalition. 2008. "UPDIG-Home." *UPDIG Universal Photographic Digital Imaging Guidlines*. Edited by Richard Anderson, Michael Stewart, David Riecks, Peter Krogh and Greg Smith, September 22. http://www.updig.org/pdfs/updig_photographers_guidelines_v40.pdf.

Warda, Jeffrey, ed. 2011. *The AIC Guide to Digital Photography and Conservation Documentation*. Washinton DC: American Institute for Conservation.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

Tess Colwell

5 Cataloging and Description Practices Informed by Rationale in a Small LAM Environment

Introduction

Brooklyn Historical Society (BHS), located in Brooklyn, New York, is a small, non-profit museum, education center, and research library. BHS was founded in 1863 as the Long Island Historical Society. Long Island is made up of four counties (Kings, Queens, Nassau, and Suffolk) that includes the borough of Brooklyn. The Long Island Historical Society's founding mission was to collect and preserve materials related to general history, especially materials related to the United States, the state of New York, and the communities of Long Island. It wasn't until the late-twentieth century that the collecting began to focus more specifically on materials within the geographic borders of Kings County. In 1985, the Long Island Historical Society changed its name to Brooklyn Historical Society, reflecting its new institutional focus. Today, BHS library staff collects, preserves, and makes available a range of material concerning Brooklyn's history, including books, archival collections, photographs, maps, artifacts, oral histories, and artwork. BHS acquires its collections primarily through donations approved by the organization's Collections Committee, which is comprised of trustees and staff. Popular research topics include housing and family research, but span a sweeping range of topics concerning Brooklyn history and culture.

The Othmer Library

The Othmer Library, BHS's landmarked library, contains one of the most comprehensive collections on Brooklyn's history and culture. Its holdings include more than 33,000 books; 1,600 archival collections; 1,200 oral history interviews; 50,000 photographs; 8,000 artifacts; 300 paintings; 2,000 maps; and thousands of drawings and works on paper dating back as far as the seventeenth century. The collections "document the commercial, residential, community, and civic development of the borough" (Brooklyn Historical Society 2015, under "About"). The library collections support scholarship and research

https://doi.org/10.1515/9783110337419-006

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

on a range of topics, connect with the community through outreach and programming, and enhance institutional projects and programming.

The Othmer Library is staffed by the Library & Archives department led by the Managing Director of Library & Archives. Under her supervision is a small library staff including one Reference Librarian, an Archivist, Exhibition Coordinator and Registrar, Digital Projects Archivist, Cataloger, and student interns. One way the Library & Archives department connects to the public is through description and digitization. For example, about ten percent of the 50,000 photographs in the collection are digitized, cataloged at the item-level, and readily available online. Roughly 35,000 photographs can also be searched using an image database available inside the library. New collections are published online on an ongoing basis. BHS aims to connect all of its collections to as many users as possible through digitization and description and connecting those records through online access tools.

Like many cultural institutions, funding plays a large part in supporting projects and operations at the BHS. This includes all departments such as: (1) Public History, (2) Education, (3) Public Programming, (4) Development, and (5) the Library & Archives. In the library, funding can play a part in cataloging and description practices. These practices primarily take an iterative approach at the BHS. All collections are processed at the minimal collection or series level. Item-level records and granular descriptions are completed if there is a rationale to justify that work. Rationale for further cataloging and description beyond the series or collection level is determined by several factors. First, when processing a collection, Library & Archives staff may deem it necessary to describe the collection materials at the item-level for user access. Second, limitations of grant projects can determine whether or not particular collection materials get digitized or receive comprehensive description. Grants sometimes allow BHS to hire additional staff to process and catalogue a large group of collection materials within a designated timeframe. Finally, rationale for enhanced description can be informed by internal projects from other departments. For example, if photographs or archival materials are needed in support of an institutional project, additional descriptions and digitization will be completed. This is not an exhaustive list, of course, and the level of description changes and evolves for a variety of reasons not mentioned here. BHS aims to make all collection materials easily accessible to its users, and makes the most of the resources available in its library and archives. In determining levels of description, rationale is used to decide which level is appropriate. Funding can play a key part in how quickly and comprehensively that is accomplished.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

Organizational and descriptive practices

Since its founding, the Library & Archives has seen a variety of institutional practices informed by non-professional and professional librarians and archivists influenced by an evolving common practice of organization, staffing resources, and funding trends and availability. Prior to 2009, archivists and librarians at BHS separated each collection by medium. If a collection included both manuscripts and photographs, the staff removed the photographs and placed them in the Photography Collection. The visual materials would be described at the item-level, and receive a different accession number from the manuscripts. All materials received file-level processing, removal of staples, photocopying of newspapers, and rehousing. In addition, there were no relationships described linking the photographs or other visual materials. This separation of media was common practice until the BHS refined its descriptive goals. Today, BHS library staff aims to reunite previously separated collections to reestablish context and enhance descriptions with improved cataloging.

In current descriptive practices at BHS, there are two distinct methods of processing. The first is *Accessioning as Processing*, informed by an article of the same name by Christine Wiedeman (2006). This approach focuses on processing a collection at the point of accessioning. Wiedeman (2006) explains, "during the accessioning process, whenever possible, we arrange and describe the materials, including the creation of the finding aid, so that they are ready for research use and never enter our backlog" (276). At BHS, if a collection is larger than two linear feet, it is processed at the series level. If the staff resources to process a large collection are unavailable, the collection goes to the processing backlog until either funding is secured or processing interns are available to assist the archivist. This is a department policy and it is determined and maintained by the Managing Director of the Library & Archives. The processing needs are considered at the time of acquisition when donors of large collections are asked by Library & Archives staff if there is money to assist in transporting, processing, rehousing, storing, and digitization – before it arrives and is taken into consideration by the Collection Committee at the time of approval or rejection.

Like many libraries, BHS has a processing backlog and limited resources, which can affect the level of description. To mitigate backlog, BHS's processing and cataloging practices are informed by a second method of processing, *More Product, Less Process* (MPLP) an influential archival approach conceived by Mark A. Greene and Dennis Meissner (2005). This approach emphasizes minimal processing, aiming to source collection information to researchers more

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

rapidly. This means that many of the collections are described minimally at the collection or series level. Using this model allows researchers to narrow down their searches significantly, based on the minimal information in the finding aid. This approach also allows BHS to process and provide more collection materials to users than if a more traditional approach was utilized, and complements the iterative approach to cataloging already in place.

The archivist develops a processing plan to determine the level of description necessary for user access. For larger collections, the processing plan is shared with the Managing Director of Library and Archives to review and develop a plan for processing and describing the collection. Decisions to develop a more granular description of collection materials are based on a combination of anecdotal evidence from interactions with library and archive users who identify a particular topic or part of the collection having particular research value, ongoing staff projects or future projects that have identified specific topics of research or exhibition use, the medium of the collection could warrant item-level description, and the Library & Archives' topical priorities established in the Collection Development Policy. For instance, topics relating to the Gowanus Canal is a popular research topic due to its status as a Superfund site, and could prompt the need for item-level description of a collection containing materials related to that topic due to high requests for researchers. Another example is if a collection contains a large amount of visual materials or photographs, particularly a collection made up solely of color negatives that are particularly difficult to look at in negative form. Photographs in such a collection might receive item-level digitization and description to provide research with basic access the objects. Because of a processing backlog and fluctuating staffing resources, describing all collections at the item-level would significantly hinder the rate at which collections are generally described and made available. The MPLP approach allows for the BHS collections to get to researchers swiftly and strategically. Item-level description is performed when there is a rationale for enhanced description and the Library & Archives staff deems it necessary for access.

Currently, BHS does not have a Digital Asset Management (DAM) system. In place of a DAM, a combination of tools is used for description. ArchivesSpace (AS) is used as a content management tool or the creation of finding aids. The description practices at the BHS currently centers on AS for the first level of description. AS is used primarily for archival collections, and includes accession, location, and minimal description formation across all collection types records are created using AS, and then exported to the New York University (NYU) Aleph/Bobcat catalog and NYU Finding Aid Portal. The NYU Finding Aid Portal brings together collections and finding aids from several repositories in New York City.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

For item-level descriptions, the collection management system PastPerfect is used for photographs, oral history collections, works on paper, fine art, and artifacts. PastPerfect offers an enhanced level of description for visual materials and allows BHS to publish images for online consumption using PastPerfect online. Images can be searched by: (1) keyword, (2) collection, (3) creator, (4) date, (5) subject headings, (6) place names, (7) people, (8) description, (9) object ID, (10) collection finding aid, and (11) date range. Images are continuously pushed online as they are described, digitized and published.

BHS's artifact collection is stored at two locations: one at BHS headquarters and the other is less than two miles away in Brooklyn. While artifacts are described at the item-level in PastPerfect, they are first given minimal description by the Exhibition Coordinator & Registrar and adopt a similarly iterative approach to archival collections of mixed media. The first goal is a basic record with an object name, photograph, and location. The second goal is to enhance the record with more description. At this time, that descriptive record is fueled by an internal or external curator's request for more information and/or to borrow the item. For instance, if a curator borrows an artifact for use in an exhibition, additional contextual information might be provided which would allow for additional description of that artifact. Similarly, if an artifact is used for a BHS exhibition or project, a researcher might provide historical context for the item, which would then be added to the record for enhanced description.

Grant funding

BHS has experienced great success through grants, which have enabled many exciting and valuable collections to reach a wider audience. It also provides a rationale for enhanced description practices beyond the series or collection level. For BHS grants, there has been a notable shift from bulk description to item-level description. For example, in 2009 BHS received a grant from Council on Library and Information Resources (CLIR) to process and make available hundreds of collections, containing maps, manuscripts, and photograph collections. This grant contributed to the processing of roughly 75 percent of BHS collection materials that were previously not accessible and mitigated a substantial processing backlog. Prior to this project, very few finding aids or guides were available to researchers. The CLIR grant approached the processing of the collections using the MPLP model of description, creating minimal finding aid records of hundreds of collections. This was a project that made significant strides in description for BHS collections.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

In 2013, a second grant from CLIR allowed for enhanced description of the finding aids, allowing for improved folder and item-level description of collections. This grant funded the processing of Brooklyn's Corporation Counsel Records that comprises documentation and maps relating to legal cases brought against the City of Brooklyn from 1843 to 1920. Without this funding, the collection would likely still be folded and in off-site storage. The second CLIR grant allowed for enhanced description of the minimally described finding aid to include description of each folder. This collection is a good example of a minimally processed collection that benefited from enhanced description for research access.

The Ortner Collection, on the other hand, is a unique example of a collection that was processed minimally, but is in need of enhanced description. This collection contains roughly 250 boxes of legal and personal documents of Evelyn and Everett Ortner. The Ortner's were instrumental in the historical preservation movement in Brooklyn, promoting the borough as a historic district worth preservation. Their focus was particularly in the Park Slope area of Brooklyn, preserving the historic brownstones that lined the streets neighborhood where they lived. They organized and raised funds for the movement, eventually leading to the designation of Park Slope as a historic district in 1973. These documents are historically relevant and valuable to researchers interested in Brooklyn housing. This collection was funded by the Documentary Heritage Program (DHP) for processing and description at the digital file level. While this would typically be more than enough for an archival manuscript collection, this collection included more than 10,000 photographs. BHS applied for another grant (not through DHP) to digitize and describe the collection's photographs at the appropriate item-level, but it was not awarded. The photographs, mostly 35mm slides, remain accessible via light tables and digital file, but will not be described at the item-level until further funding is found.

Grants are immensely valuable to a small cultural institution like BHS and can provide the rationale for item-level description of collection materials. Additional funding allows the institution to dive into a specific project and provide enhanced access for researchers. The institution's ability to construct item-level descriptions for collections can be dependent on additional funding from grants. Grants support a range of projects at BHS including cataloging, digitization, preservation, conservation, and digital humanities projects. In particular, grants support the digitization and item-level descriptions for many photographic materials at BHS. For example, the John D. Morrell Photograph Collection is fully digitized and cataloged at the item-level, thanks to a grant from Gerry Charitable Trust in 2009. This collection comprises more than 2,000 black-and-white and color photographs, taken between 1958 and 1963. This col-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

lection was identified as one that would benefit from item-level digitization through anecdotal evidence gleaned from reading room researchers because architectural history is one of the most popular research subjects at BHS, and the photographs in this collection depict street scenes from nearly every neighborhood in Brooklyn accompanied by descriptive metadata from the photographer. This collection is popular with many different types of researchers including members of the public interested in learning about their homes and scholars interested in studying architectural styles throughout history. The photographer, John D. Morrell, was a former librarian at BHS and indexed all of the photographs with detailed addresses, street names, and neighborhoods depicted in the Brooklyn street scenes. That information has been transformed into rich metadata for improved searchability. That project has continued to benefit researchers and the institution. One benefit of a grant-funded project is that the entire collection can be digitized and cataloged in detail, which improves access to collections having high research value. It also allows for the institution to hire and maintain highly qualified staff to support the project and the institutional mission.

Similarly, BHS recently received another grant from Gerry Charitable Trust to digitize and describe eight scrapbooks, and more than 2,000 photographs compiled by amateur photographer Eugene L. Armbruster in the early twentieth century. This project is extremely valuable for the breadth and content of the scrapbooks, which offer a candid street photography view of Brooklyn. The project includes digitizing and cataloging thousands of photographic records at the item-level and hiring an additional staff personnel and student interns to complete the project. At the completion of the project (an estimated two years from start to finish), thousands of photographs will be digitized, described, and made available for research use.

This project involves the construction of item-level descriptions of the photographs contained within the scrapbook which includes the interpretation of caption details provided by Armbruster into standardized metadata for online discovery. The BHS utilizes a modified (by BHS) Dublin Core (DC) metadata schema incorporating Library of Congress Subject Headings, Getty Art & Architecture Thesaurus (AAT), and local geographic headings specific to Brooklyn and Long Island.

The downside to a description model influenced by grant funding is that sometimes only select collections (or parts of collections) are fully described at this level, while other collections without grant funding often have less detailed descriptions. In some cases, less access means the collection materials are only accessible in-person rather than online. For instance, the Armbruster Collection includes several loose photographs, not included in the album pages. These

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

photographs will not be included in the digitization project because they are beyond the scope of the grant. Still, the Gerry Charitable Trust grant enabled BHS to digitize and describe the bulk of the Armbruster Collection, which had limited access to research prior to this project.

When it comes to grant applications, the library staff has to make choices in determining what collections or project should take priority with funding. Evaluation criteria for grant projects include the consideration of exhibition value, research value, condition of the collection, and rarity. For digitization projects, copyright is also taken into consideration. For instance, if the collection is owned by BHS or if the work is in the public domain, it will take precedence over work that has copyright restrictions. Every funder has differing criteria and stipulations for funding, and each proposal considers how BHS collections can fit into the criteria to fund collection projects. An increasing amount of funders require that materials submitted for digitization projects must have a known copyright holder or rights must be transferred to the digitizing repository. For example, the CLIR Digitizing Hidden Collections program clearly dictates that one eligibility requirement is copyright: "Materials that are in the public domain in analog form must continue to be in the public domain once they have been digitized. CLIR strongly encourages grant recipients to share digitized collections as public domain resources or with creative commons licenses, as appropriate" (CLIR 2016, under "Eligibility" on the "For Applicants" page). Grant funding sometimes forces a rationale as opposed to providing one. The unique criteria sometimes mean a collection might receive more or less description than the Library & Archives staff would do otherwise.

BHS library department supports public research efforts, but also institutional projects. When another BHS department is working on an exhibition or educational curriculum, collection materials may be used in support of that project. This provides an additional rationale for enhanced description for the collection items used in conjunction with another project. A particular collection or collection item receives additional description in support of its use in exhibitions, educational curricula, or other institutional projects. For example, the BHS education department developed a curriculum on the Civil Rights Movement. For that project, BHS digitized a flyer and photographs from the Arni Goldwag Brooklyn Congress on Racial Equality (CORE) Collection, 1943–2007. Collection materials were digitized, described at the item-level, and added to the Archived Digital Assets server for preservation and storage. This is beneficial to researchers because many of BHS's institutional projects have an online component. Researchers can find collection materials through those project websites and through the online gallery. The downside to enhancing select materials for a project is that only select items (not entire collections) receive en-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

hanced item-level descriptions. In the case of Brooklyn CORE, many of the visual items in that collection were digitized and described for use in this project; however, the entire Brooklyn CORE collection was not described at the item-level because they were out of scope for the specific project.

BHS exhibitions also provide a rationale for enhanced description for collection materials. An upcoming long-term exhibition entitled *Waterfront* will include hundreds of photographs and visual materials from the BHS collections. The exhibition's location at Brooklyn Historical Society DUMBO (or DUMBO, which is short for Down Under the Manhattan Bridge Overpass), inside the recently redeveloped nineteenth-century Empire Stores warehouse on the Brooklyn waterfront, provides new opportunities for enriched description of collection materials.

Two digital installations, *Water Log* and *Visitor Vistas*, will allow visitors to engage with BHS collection materials in new ways. Visitor Vistas is a touch screen digital interactive experience that allows visitors to insert themselves into historic images of Brooklyn's waterfront held within BHS collections. Ranging from seventeenth to twenty-first century paintings, works on paper, and photographs, BHS image collections depict an array of waterfronts, from pre-industrial landscapes to ghostly post-industrial neighborhoods. Water Log introduces visitors to more than 20,000 years of Brooklyn's waterfront history through floor to ceiling projections of BHS collections. The installation transports visitors into a series of historical moments, pairing renderings of the landscape with the sights and sounds of the natural and working waterfront.

To support this project, Library & Archives staff digitized hundreds of photographs, maps, works on paper, and more, while updating and providing detailed descriptions of the materials. The *Waterfront* project bolstered institutional rationale for greater description efforts within the Library & Archives to serve both institutional and public needs. The Education and Public History departments, for example, rely on the materials and their associated description for research and teaching, while the public engages with the collections when visiting the library or pursuing personal or professional research interests.

Digitization is an important component of item-level description because it ensures the preservation and discovery of items today and in the future. It is not a rationale in and of itself for description, but rather a method. For photographic collections and visual materials, in particular, digitization and item-level description go hand in hand. Institutional projects and grants provide the rationale to digitize and describe collection items. Researchers and library users, however, benefit from these projects because more collection items are published online and available for research and scholarship.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

Benefits of digitization

In a digital world, digitization is the norm or the basic expectation of most researchers and is necessary for most repositories in order to connect researchers with collections of high research value. Converting analog collections into a digital format has many benefits. The most obvious benefit is reaching a wider user-base. This includes connecting collection materials to users anywhere that the Internet can be accessed, without having to physically travel to the archive. Connecting collection materials to a larger audience is also a way to promote the collections, to gain new visitors, and encourage the continued collecting of the institution. Reaching a wider audience also opens up more opportunities for collaboration and context of the collections. At BHS, users can access a portion of digitized photographs through an online image database. There is a link to comment on the image directly through the image page. On several occasions, researchers have utilized the response link to offer additional contextual information on the image, which provides another opportunity for record enhancement.

Digitizing also protects the analog format of collections, particularly if they are fragile, fading, or in an outdated media type. Having a digital format of the item protects those collection materials from additional wear and tear from handling. It also ensures that the material can be used for future generations, even beyond the life of the item itself. For example, BHS's oral history collection contains interviews on a variety of media types, some of which are obsolete. The digitization and description of these oral history collections ensure that these interviews can be accessed and used in scholarly pursuits for years to come.

Lastly, digitization provides endless opportunities for the institution and for collaboration. Having detailed description and digitized collections enables every department within BHS to use the digitized collection materials to promote the institution's mission and services. Having digital collections published online also allows for collaboration with other institutions and organizations to reach a wider audience by sharing collections. For example, BHS collaborates with the Digital Public Library of America (DPLA) to push out more records. That collaboration and sharing of records is an advantageous method of exposure that fuels BHS's cataloging.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

Conclusion

Like any library or special collection, BHS strives to provide access to all of its collection material. In an ideal world, BHS would describe all collection items in detail at the granular level. Limited resources and reliance on grant support means that cataloging decisions stem in large part from what BHS can accomplish with the resources it has available at the time. Funding plays a significant role in cataloging practices and the level of description for each collection.

Description and cataloging practices at BHS takes an iterative approach. Following both the MPLP and Accessioning as Processing Models, all collections are described at the series or collection level, to allow researchers to access the collections and narrow their areas of interest. This also helps with mitigating processing backlog and getting collection materials to researches more rapidly. Collections are described at a more granular level when there is a rationale to do so. Rationale for item-level description includes several factors. First, the Library & Archives deems that item-level description is imperative for access. Second, BHS receives a grant that enables the detailed description of an entire collection, or significant portion of the collection. Last, institutional projects and exhibitions rely on collection materials. These projects provide an additional rationale to create or enhance item-level records and digitize materials. There are exceptions to these descriptive practices, and this list is not exhaustive. In this model, rationale leads to description, then leads to multiple levels of access for researchers and staff, which is an ongoing process. BHS is continuing to refine its cataloging and descriptive practices to ensure that collection materials are reaching as many users as possible, as quickly as possible.

Relying on rationale for granular description means that many collections will not receive enhanced description. With limited resources, rationale ensures that collection materials are not described randomly, but in support of scholarship and research. Many researchers, students, and casual learners have benefited from the collections available at BHS and online. Collections lose value if they are not accessed and used for scholarly and personal pursuits. Through rationale, the goal is to establish structure in allocating the limited resources available to collections with particular historical or cultural relevance and to reach the widest user group possible.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

References

Brooklyn Historical Society. 2011. "Guide to the Arnie Goldwag Brooklyn Congress of Racial Equality (CORE) Collection ARC.002." *Brooklyn Historical Society*. http://dlib.nyu.edu/findingaids/html/bhs/arc_002_goldwag_core/arc_002_goldwag_core.html.

Brooklyn Historical Society. 2015. "Othmer Library." *Brooklyn Historical Society*. http://brooklynhistory.org/library/about.html.

Brooklyn Historical Society. 2016. "Guide to the Everett and Evelyn Ortner papers and photographs ARC.306." *Brooklyn Historical Society*. http://dlib.nyu.edu/findingaids/html/bhs/arc_306_ortner/arc_306_ortner.html.

Brooklyn Historical Society. 2012. "Guide to John D. Morrell Photograph Collection ARC.005." *Brooklyn Historical Society*. http://dlib.nyu.edu/findingaids/html/bhs/arc_005_morrell/arc_005_morrell.html.

Council on Library and Information Resources. 2014. "For Applicants." *Council on Library and Information Resources*. https://www.clir.org/hiddencollections/applicants.

Greene, Mark A., and Dennis E. Meissner. 2005. "More Product, Less Process: Revamping Traditional Archival Processing." *American Archivist* 68 (2): 208–263. doi:10.17723/aarc.68.2.c741823776k65863.

Weideman, Christine. 2006. "Accessioning as Processing." *American Archivist* 69 (2): 274–283. doi:10.17723/aarc.69.2.g270566u745j3815.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

Jane Zhang

6 Digital Archival Representation: Approaches and Challenges

Introduction

Archival representation treats the context and content of an archival collection as an organic whole and integrates them into one hierarchical structure. The movement of digitizing hidden collections turns archival materials into digital objects and provides them with granular, multi-faceted, item-level metadata. At the same time, born-digital archival materials start to be accessioned into archival collections. How will archival context and digital object metadata coalesce in digital archival representation? This chapter reviews digital archival representation practices and discusses approaches and challenges in maintaining archival context and digital content accessibility and interoperability in digital archives.

Literature review

Traditionally, provenance-based hierarchical archival finding aids are considered sufficient to represent and gain subject access to archives due to the inadequacy of content indexing of archival materials (Lytle 1980). Encoded Archival Description (EAD), developed in the 1990s, makes finding aids available for use in a networked environment and allows the online search of collection information in finding aids within and across repositories. The unique characteristics of markup language provide an enabling tool for multi-level presentation of archival descriptive information from collection, series, folder, to individual items. However, using EAD as a platform for digital collection presentation may be insufficient if traditional folder or item captions from finding aids are used for searching and identification of the digital content. Very little guidance is available for archivists to describe digitized items from archival collections (Hensen 2001).

The library community developed item-level metadata creation standards used for describing digital objects. In her survey of literature published from 2000–2005 on library digitization projects, Lopatin (2006) listed major metadata standards adopted to describe electronic resources and digital objects, in-

https://doi.org/10.1515/9783110337419-007

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

cluding (1) Dublin Core (DC), one of the most widely accepted metadata standards, (2) library cataloging MAchine-Readable Cataloging (MARC) format, updated to accommodate the cataloging of electronic resources, and (3) two new metadata schemes developed primarily to accommodate digital resources: the Metadata Object and Description Schema (MODS) and the Metadata Encoding and Transmission Standards (METS). The literature she reviewed also covered several controlled vocabularies used to provide access points to digital resources, including (1) the Library of Congress Subject Headings (LCSH), (2) the Getty Art & Architecture Thesaurus (AAT), (3) the Getty Thesaurus of Geographic Names (TGN), and (4) the Library of Congress Thesaurus for Graphic Materials (TGM) (Lopatin 2006).

Early adopters of metadata standards for digital projects were reported in a survey of digital projects funded by the Institute of Museum and Library Services (IMLS) National Leadership Grant program from 1998–2002. The survey found that eighty-six percent of the respondents used item-level metadata. Among them, fifty-six percent used DC, thirty-three percent used MARC, and thirty-nine percent used locally developed standards. In terms of controlled vocabularies, eight-four percent used controlled vocabularies for topical terms for subject access. Among them, seventy-three percent used LCSH and twenty-seven percent used the TGM (Cole and Shreeves 2004). Interoperability was another topic covered in the digital project literature. Tools used to improve the interoperability of digital collections include the ANSI/NISO Z39.50 protocol, the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH), and a variety of crosswalks. Researchers also acknowledged difficulties in overcoming interoperability challenges due to semantic differences and diversity of controlled vocabularies and metadata schemes used in digital projects (Lopatin 2006).

In August 2007, more than 2,000 professionals responsible for special collections attended a one-day event co-sponsored by Research Libraries Group (RLG) and the Society of American Archivists (SAA). The event called for increasing the scale of digitization of special collections by applying archival collection-level control and moving away from obsessive item-level description (Erway and Schaffner 2007). The proposal reflected the trend of digitization of large segments or entire collections of archival materials that started in the 2000s. These large-scale archival digitization projects depart from item-level metadata and use folder-level description in the finding aids created during the processing (Miller 2013).

The debate about different practices in organizing and describing digital archival materials makes further research necessary to investigate different approaches and challenges relating to digital archival representation. This chapter

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

discusses three digital archival collection sites, each representing a unique digital representation approach. The three sites are (1) the Archives of American Art's (AAA) Terra Foundation Center for Digital Collections, (2) Washington State Digital Archives, and (3) Digital Public Library of America. The chapter concludes by discussing different levels of digital archival representation based on the data drawn from the three digital collection sites.

Archives of American Art's (AAA) Terra Foundation Center for Digital Collections

The Archives of American Art's (AAA) Terra Foundation Center for Digital Collections is a digital repository that contains more than 125 archival collections "scanned and posted online in their entirety" since 2005 (AAA 2016. Terra Foundation Center for Digital Collections, par. 1). This substantial achievement is closely associated with the Archives' innovative large-scale digitization program that can "trace its beginnings to the Archives founding mission to provide access to entire collections using microfilm and distributed via interlibrary loan" (AAA 2016. Terra Foundation Center for Digital Collections, under "About").

Founded in 1954, the Archives of American Art was initially established as a central repository and distributor of microfilmed archival collections (mostly papers of artists, dealers, critics, historians and other art world figures) to support research in the field of American art history. The first 109 reels of microfilm were produced from the holdings in libraries, museums and historical societies in Philadelphia. After that the microfilming expanded to New York City and then across the country. At the same time, the Archives started to acquire original papers from individual donors and then microfilm them to facilitate distribution through interlibrary loan. After half a century's acquisition and microfilming, by 2004 approximately one-third of the Archives' collections had been microfilmed, totaling over 11,000 unique reels (Weiss 2010).

The mission-driven approach of microfilming to provide access to the entire collections since its founding in the mid-1950s, together with the adoption of Encoded Archival Description (EAD) in the late 1990s, laid a solid foundation for the Archives' innovative large-scale digitization program initiated in the mid-2000s. In 2005 and 2011, the Terra Foundation awarded the Archives of American Art two multi-million dollar grants to support two five-year digitization projects. The Archives, building on its half-a-century entire collection mi-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

crofilming experience, launched a substantial large-scale digitization program. The "unprecedented and innovative presentation dubbed *Collections Online*" (Weiss 2010, 104) was made public in 2006 with an inaugural release of ten fully digitized collections. By 2010, eighty-two collections were scanned and added to the *Collections Online*. As of May 2016, the list of digitized collections is close to 160. The majority of them are personal papers but records of art galleries, association, and art museum are also included.

There are three key components of the Archives of American Art (AAA's) innovative archival approach to large-scale digitization that empower the organization and representation of AAA's digital collections: (1) fully digitized archival collections that maintain the same order as their physical arrangement, (2) eXtensible Markup Language (XML) based Encoded Archival Description (EAD) finding aids that break down the archival description into a logical hierarchical structure, and (3) carefully designed technical infrastructures and workflows that combine digital files and archival description and produce an effective online display of the digitized collections on the *Collections Online* website (Weiss 2010).

By the first decade of the twenty-first century when the archival profession started to recognize the value of mass digitization, the Archives of American Art had microfilmed entire archival collections for several decades and begun to look into a logical transition from microfilming to digitizing entire collections in order to make them accessible online. The Archives discovered that "the measures" outlined in mass digitization could easily be adopted to replace what they used for microfilming and help them update "from an access program based on a dying and unpopular microfilm technology to a modern, web-based platform" (Weiss 2010, 111). The early endeavor of digitizing microfilm started with a project "to microfilm all 109 feet of the Downtown Gallery records" in 1999 that "ultimately succeeded in presenting the Downtown Gallery records online in 2002" (Weiss 2010, 111–112). The project software developer's choice of a flexible scripting language made it possible for the Archives to easily adopt "the EAD-XML solution for presenting fully digitized collections" (Weiss 2010, 112). As Weiss (2010) explained, the 1999–2002 project:

> unintentionally brought together what would become the three strategic components of *Collections Online*: an XML-based EAD finding aid providing structured data describing the collection overall and its groupings of series and subseries, and their respective containers and folders; digital files stored in a directory structure corresponding to the EAD finding aid and named in such a way to define sequence and arrangement; and web programming to render the structured data useful in creating the all-important online presentation. (111–112)

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

To accommodate the multi-level finding aid data, the Archives expanded the existing database by breaking into "high level, series and subseries level, and container level tables" and creating "a separate digital resource table to manage the linking of the folder data to the digital files" (Weiss 2010, 114). The multi-level table structure accommodates ingest of the complete XML finding aid files, and allows "the web programming to take full advantage of the encoding to deliver the finding aid to the web in a presentation reflecting the collection's arrangement" (Weiss 2010, 114). The digital files following file-naming conventions are stored in the directory structure, which "provides the necessary information to generate a unique ID in the digital resource table for each image, identifying its correct box and folder designation within the collection, as well as its sequence within each folder" (Weiss 2010, 114).

Using EAD finding aids as the source of descriptive and structural metadata for digital files, archival collections are scanned in the same order as their physical arrangement regardless of format. Because the file folder is the lowest level of description, access to the digital files is done at the folder level, not at the document level. The lack of item-level access in the digitized archival collections had made it necessary for the Archives of American Art to create a separate system to host and provide access to individually digitized items created in response to requests by researchers and staff. These items (currently more than 12,000) are individually catalogued and accessible online via the Archives of American Art Image and Media Gallery page. The item-level metadata contains creator, description, subjects, physical details, forms-part-of, rights statement, subjects, and digital ID. The forms-part-of field links the item to the source collection, but the lack of item-level metadata in the archival collection description makes it impossible to pinpoint the item to its appropriate placement in the collection.

Washington State Digital Archives

State archives were well established in the twentieth century, but they are faced with the tasks of providing access to the oldest records in their custody, some of which predate statehood or even the creation of the United States, as well as to their most contemporary records in their holdings (Walch 1997; Turnbaugh 1997). Similar to any other archives, state archives hold collections that are organized by provenance (e.g., governors' records, state agencies and departments records) or accumulated based on media or genres (e.g., photographs, audio-visual, sound recordings, maps, architectural drawings). Unique in the

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

state archival holdings are large groups of records generated to support specific functions state governments perform (e.g., birth, marriage and death records, land and property records, census and naturalization records, state and county court records). The unique characteristics of state archival records made it possible for them to launch some pioneering initiatives in building digital archival collections.

The Washington State Archives hosted its "much-anticipated grand opening" on October 4, 2014 to celebrate the official launching of "America's first state government digital archives" (Washington State Archives – Digital Archives 2014, first par.). The digital archive system went online that day with the electronic marriage records from three pilot counties and the historic census and naturalization records database from the State Archives and State Library. The Digital Archives is also claimed to be "the world's first built-from-the-ground-up digital archive" (Washington State Archives – Digital Archives 2016, first par.) because its state-of-the-art facility, located in Cheney, WA which opened in 2004, "was designed from the ground up to host this technically complex program" (Washington State Archives – Digital Archives 2016, under "Digital Archives Background").

The specific objectives of the Digital Archives are to make the historical electronic records of State and Local government easily accessible to the public. The web interface and database were custom designed specifically for the Digital Archives to hold the unique and important electronic and digitized records found throughout the state, and to provide simple, straight forward access to researchers. Microsoft Consulting Services, in partnership with the Digital Archives, designed a customized solution for ingesting and archiving electronic records coming into the Digital Archives for long-term storage and retrieval. The solution uses Microsoft SQL Server and Windows Workflow Foundation as the core technology to connect hundreds of state and local government offices with the Digital Archives so that legal and historical records can be automatically and electronically transmitted and archived with no human involvement of any kind (Washington State Archives – Digital Archives 2016, under "Partners").

Online collections within the Washington State Digital Archives are organized into three levels. At the top level, "records that are related as a result of being created, received or used in the same activity" are grouped into various record series (Washington State Archives – Digital Archives 2016, under "Frequently Asked Questions"). Currently there are more than thirty record series, mostly based on records generated from state functional activities. At the second level, each record series breaks down into geographic or organizational divisions. At the third level, documents in each record series are indexed and made searchable through field specific metadata. Due to the different character-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

istics of each record series, different metadata schemas are applied to each of them to facilitate specific searches. For example, birth records are indexed by mother/father/child's name, gender and year; death records are indexed by person's name and year; ordinance records are indexed by ordinance number, keyword, and year; and real property records are indexed by parcel number, house number, street name, and legal property description. The function-based document-specific descriptive approach has a direct impact on search strategies developed for online collections in the Washington State Digital Archives.

The Digital Archives also provides several search options to help users find what they are looking for from the digital collections. Each search is suited for different needs. The *People Search* for example, allows information users to search all records in the digital collections that have name-based index fields. *Keyword Search* allows one to search on all other records in the digital collections that are not name-based records. The keyword field allows the information user to enter one or more terms to locate records that contain those words. Both *People Search* and *Keyword Search* are designed based on the description strategy that documents in the Digital Archives are indexed at the basic level by person/party's name, or by keyword for non-name based records. They are made searchable across all collections, and at the same time, information users can also select a particular record series from the drop down list to narrow the search to one record series (e.g., marriage, census, executive orders). Built on their individually indexed record series, the Digital Archives offers a third search option, *Detailed Search*, which allows a user to select one record series and search on additional fields that are pertinent only to that record series (such as bride and groom for marriage records). Researchers can also limit their searches to counties or time frames within specific record series to obtain specific search results.

The Washington State Digital Archives breaks down its online collections into individually indexed record series to help achieve more granular search results. In order to allow the information user to access information at such granularity, the database needs to supply each individual record with the appropriate item-level metadata. To expand the supply of quality metadata, the Digital Archives lunches Scribe, a crowd-sourcing tool that allows registered users to select record series, view individual documents, and supply metadata into a pre-defined template (Washington State Archives – Digital Archives 2013). Due to similarity of records and service needs, other states have turned to the Washington State Digital Archives for inspiration and guidance. Building on the Washington model has also helped other state digital archives preserve and provide access to their digitized archival collections (Washington State Archives – Digital Archives 2016, under "Background and History").

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

Digital Public Library of America

Despite the Boston Marathon tragedy that occurred in front of the Boston Public Library where the Digital Public Library of America (DPLA) gala launch had been planned, the new DPLA site was launched as planned at noon ET on Thursday, April 18, 2013 (DPLA 2013, under "A Message from Executive Director Dan Cohen"). The launch of DPLA was a major breakthrough toward realizing the vision of a national digital library, a dream that "has been circulating among librarians, scholars, educators, and private industry representatives since the early 1990s" (DPLA 2016, under "History").

To function as a national digital library, the DPLA operates on a national network using a Hub model. The Content Hubs, made of large libraries, museums, archives, or other digital repositories, provide unique metadata records associated with their digital objects to the DPLA. The Service Hubs, made of state, regional or other collaborations, bring together digital objects from libraries, archives, museums, and other cultural heritage institutions and provide the DPLA with their partners' unique metadata records for digital objects through aggregated metadata feeds. In this way, digitized and born-digital content from across the country are brought into a single access point for end users as well as an open platform for developers. John Palfrey, who once served as director of the Harvard University's Berkman Center and member of the DPLA Board of Directors, thus explained how the structure of these Hubs offers a preview of what the DPLA would accomplish on a national level: "Each of the Service Hubs collects content from its region, assisting the libraries and archives responsible for that content in the process. The DPLA will help to aggregate the metadata and the materials and in turn connect users to the content across these many regional and institutional digital libraries" (Palfrey 2013, under "Distributed Effort").

The DPLA has made steady progress after its successful planning and launch toward achieving its ambitious goal of bringing together the riches of America's libraries, archives, museums, and cultural heritage sites, and making them freely available to students, teachers, researchers, and the general public. At its first anniversary in April 2014, the digital item and institutional coverage of the DPLA increased significantly from 2.4 million items and 500 institutions to more than 7 million digitized cultural heritage items from 1,200 contributing institutions across the United States (DPLA 2015, under "Strategic Plan"). The top priority set in the DPLA's 2015–2017 strategic plan is to complete its national network of Service Hubs so as to draw materials from every state into the PDLA (DPLA 2015, under "Strategic Plan"). As of May, 2016, there were twenty one Service Hubs listed on the DPLA hubs page, among them eight new Service

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

Hubs that were added in 2015, covering states of Florida, Maryland, Tennessee, Illinois, Main, Michigan, Pennsylvanian, and Wisconsin.

The approach adopted by the DPLA in building this national digital library platform is to aggregate metadata collaboratively contributed by regional hubs. The foundation of this approach is a large number of digitized or born-digital objects built into digital collections by local and regional libraries, archives, and museums. Many of the digitized collections constructed by libraries, archives, museums, and other cultural heritage organizations in the past two decades have been supplied with standard or locally designed item-level metadata. For example, CONTENTdm is Dublin Core (DC) based, supports the Metadata Encoding and Transmission Standards (METS) schema, and uses major controlled vocabularies or thesauri such as Art & Architecture Thesaurus (AAT), Medical Subject Headings (MeSH), Thesaurus of Geographic Names (TGN), and Union List of Artist Names (ULAN). dSpace uses Dublin Core by default, but multiple schemas can be held in the metadata registry, and controlled vocabularies are also associated with correspondent DC metadata fields. Fedora is designed to manage all types of content and associated metadata in any format (e. g. RDF, XML) and supports ingest of objects in a Fedora-specific extension of METS. Dublin Core metadata submitted or generated in Fedora are encoded within METS. There are many other user communities who use their own locally developed metadata schemas to meet the description and management needs of their digital repositories.

The key technical design of the DPLA is its Metadata Application Profile (MAP), which serves as the basis for how metadata is structured and validated in DPLA. The DPLA MAP enables the integration of metadata created and shared by the Hubs in a variety of metadata standards. The robust and flexible structure of the DPLA MAP is achieved through describing entities (such as the resource being described, the record describing it, the person who created it, and the place it is located) independently and then creating relationships between them. This structure allows many metadata standards (such as DC, METS, MODS, MARC XML) to interface with the DPLA Metadata Application Profile and also provides for crosswalks from some local metadata application profiles. DPLA also incorporates enrichment services to check name and place authorities in order to enhance the original metadata records from partner institutions with additional authority metadata, or to standardize date formats (e.g., yyyy-mm-dd) and other clean-up tasks (DPLA 2015, under "An Introduction to the DPLA Metadata Model").

The DPLA, through its leadership and community engagement, continues to develop new initiatives for achieving its ambitious goal of building a national digital library platform. Although major content providers can work with the

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

DPLA in a one-to-one relationship to provide their metadata records, small and medium sized institutions rely on regional Service Hubs that bring together digital objects from local partners and provide DPLA with their metadata records. In preparation to become a DPLA Service Hub for the state of Pennsylvania, the newly-formed Pennsylvania Digital Collections Project recently developed a Hydra-based aggregator to begin ingesting digital content from its partners. The tool will be used for data review, normalization and remediation for the contributing institutions and the implementation team. The DPLA's Archival Description Working Group has been working on the unique challenges of archival materials presented for aggregation in the DPLA's item-level, metadata-oriented digital collections and plans to release its recommendations in summer 2016.

Levels of representation in digital archives

The three cases of digital collections presented here represent three different approaches to organizing and describing digital archival materials. In the first case of the Archives of American Art's Terra Foundation Center for Digital Collections, collections are grouped by provenance and the lowest description is at the folder level. In the second case of the Washington State Digital Archives, collections are grouped by document function and each functional group has its specific metadata fields. In the third case of the Digital Public Library of America, collections at the local repositories are described at the item-level by following Dublin Core metadata standards, then metadata consisting of various standards are integrated into a single metadata model as digital objects are ingested to the DPLA. Three different approaches reveal various levels of archival representation in terms of context and digital content accessibility.

The Archives of American Art (AAA) approach follows the archival hierarchical structure closely, and embeds digital objects (with limited descriptive metadata) into the folder level description. This approach highlights the archival context by following the principles of provenance and original order and multi-level description. However, it does not supply item-level metadata. Items in one folder are digitized and displayed as a series of images one after the other. The approach facilitates content browsing from collection summary to series description then to folder level description. It also provides keyword search for collection/series/folder level description. As no item-level metadata are associated with digital objects, accessibility and interoperability are very limited at the item-level in this approach.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

The Washington State Digital Archives (WSDA) approach organizes digital objects into document groups. Each document type supports a certain business function and shares similar documentary features and therefore can be represented through a unique metadata schema. The Archives is formed by a series of document type-based collections; each is searchable through its document specific metadata field. Cross-collection search is also provided through keyword or common field metadata. The document group approach maintains a certain level of archival context by keeping together records of a functional group and preserving their documentary relationships. It also maintains a certain level of accessibility and interoperability through common metadata schema within document groups. Although multiple collection search is limited as a result of diverse metadata schemas, it offers common field metadata search in addition to keyword search.

The Digital Public Library of America (DPLA) approach provides a standard metadata application profile to all digital objects in collections. It is a generic approach as all items are described by centrally designed metadata entities. As a result, items in digital collections can be accessible through commonly shared entities and multiple access points. The generic metadata approach provides limited archival context. However, it enjoys a higher level of accessibility and interoperability as a result of standard metadata description applied at the item-level description. The three cases of digital collections reveal three levels of representation of archival context and digital content in digital archives, as displayed in Table 6.1.

Tab. 6.1: Three Levels of Representation of Archival Content in the Digital Archives.

Digital Collections	Representation Approaches	Archival Context Representation	Digital Content Representation
AAA	Archival Approach	High	Low
WSDA	Document Group Approach	Middle	Middle
DPLA	Metadata Approach	Low	High

"Archival Approach," as shown in the AAA example, manages to retain archival context of provenance and original order, but a lack of granular metadata makes it impossible to search for detailed information within and across collections. As Table 6.1 demonstrates, it retains a high level of archival context representation but a low level of digital content accessibility and interoperability. On the other hand, "Metadata Approach" as shown in the DPLA example, takes advantage of item-level metadata assigned to digital objects to facilitate search and analysis, but fails to incorporate archival collection description where ar-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

chival context traditionally resides. As a result, it retains a high level of content accessibility and interoperability but a low level of archival context representation. The middle way "Document Group Approach," as shown in the Washington State Digital Archive (WSDA) example, manages to retain the functional context of records and relations with creators, and its series-based metadata provide field-specific search within collections and basic cross-collection metadata search. In other words, it retains a certain level of archival context but not as robust as that of the archival approach; and a certain level of digital content accessibility and interoperability, but not as robust as that of the metadata approach.

Conclusion

The reverse order revealed in the analysis reflects a challenging task for archivists in an effort to bring archival context and digital content into an organic whole in digital archival representation. If archival context representation needs to be achieved at the expense of digital content representation or vice versa, the mission of digital archival representation is partially fulfilled. Digitization of archival materials for online access has helped to increase the digital presence of archival collections, but opportunities come along with challenges as archivists would like to make sure that digital archives are not decontextualized when exposed online. The effort to bring archival context and digital content together in digital archival representation may still fall short, but the mission remains important and worthwhile for archivists to pursue.

References

Archives of American Art (AAA). 2016. "Terra Foundation Center for Digital Collections." http://www.aaa.si.edu/collections/online.

Cole, Timothy and Sarah Shreeves. 2004. "Search and Discovery across Collections: The IMLS Digital Collections and Content Project." *Library Hi Tech* 22 (3): 307–22. doi:10.1108/07378830410560107.

Digital Public Library of America (DPLA). 2013. "A Message from Executive Director Dan Cohen." https://dp.la/info/get-involved/events/launch/.

———. 2015. "An Introduction to the DPLA Metadata Model." http://dp.la/info/wp-content/uploads/2015/03/Intro_to_DPLA_metadata_model.pdf.

———. 2015. "Strategic Plan: 2015 through 2017." https://dp.la/info/about/strategic-plan/.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

——. 2016. "History." https://dp.la/info/about/history/.
Erway, Ricky and Jennifer Schaffner. 2007. "Shifting Gears: Gearing Up to Get into the Flow, Second Edition." *OCLC Research*. doi:10.25333/C3159X.
Hensen, Steven L. 2001. "Archival Cataloging and the Internet: The Implications and Impact of EAD." *Journal of Internet Cataloging* 4 (3–4): 75–95. doi:10.1300/J141v04n03_05.
Lopatin, Laurie. 2006. "Library Digitization Projects, Issues and Guidelines: A Survey of the Literature." *Library Hi Tech* 24 (2): 273–289. doi:10.1108/07378830610669637.
Lytle, Richard H. 1980. "Intellectual Access to Archives; I. Provenance and Content Indexing Methods of Subject Retrieval." *The American Archivist* 43 (1): 64–75. http://www.jstor.org/stable/40292268.
Lytle, Richard H., and W. Theodore Dürr. 1980. "Intellectual Access to Archives: II. Report of an Experiment Comparing Provenance and Content Indexing Methods of Subject Retrieval." *The American Archivist* 43 (2): 191–207. http://www.jstor.org/stable/40292291.
Miller, Larisa. 2013. "All Text Considered: A Perspective on Mass Digitizing and Archival Processing." *The American Archivist* 76 (2): 521–541. doi:10.17723/aarc.76.2.6q005254035w2076.
Palfrey, John. 2013. "What is the DPLA?" *Library Journal* 138 (7). http://lj.libraryjournal.com/2013/04/future-of-libraries/whats-is-the-dpla/#_.
Scribe. 2016. "Documentation: Introduction and Index." https://scribenet.com/wfdw/docs/.
Turnbaugh, R. C. 1997. "Information Technology, Records, and State Archives." *The American Archivist* 60(2): 184-200.
Walch, V. I. 1997. "State Archives in 1997: Diverse Conditions, Common Directions." *The American Archivist* 60 (2): 132–151.
——. 2014. "Treasures of the Archives: Archives Open House in Cheney!" https://www.digitalarchives.wa.gov/News/View/108.
——. 2016. https://www.digitalarchives.wa.gov/
Weiss, K. B. 2010. "An Archival Approach to Digitization and Web Accessibility at the Archives of American Art." *Collections: A Journal for Museum and Archives Professionals* 6 (3): 103–124.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:07 AM

Mike Jones

7 Linking Items, Connecting Content: The Donald Thomson Collection

Introduction

Contemporary collection documentation is the result of complex histories and processes, including the effects of resourcing, technological change, and evolving organizational approaches to dissemination and display. This chapter will use the Donald Thomson Collection as a case study to explore these ideas in more detail, demonstrating that, rather than a clear linear progression from internal analog practices to digital accessibility, the recent history of collection organization and documentation is more complex. Examining the past also suggests ways in which contemporary practice could be reshaped to further support users and communities in an increasingly digital world.

Museums Victoria

Museums contain layers of intersecting collections. In Australia's south eastern state of Victoria, Museums Victoria (established in Melbourne in 1854) is responsible for the State Heritage Collection, which in turn contains a multitude of collections drawn together based on people, organizations, biological classifications, subjects, formats, geographic locations, time periods and more. Former Director of the University of Queensland Anthropology Museum Leonn Satterthwait suggests that, given the varied ways in which the term is applied, the key feature of collections is "that they be regarded as such," with "individual elements connected by webs of socially engendered meanings" (Satterthwait 2008, 48). Such a definition is useful, in that it recognizes the diversity of the concept of 'the collection' and embraces the idea that such aggregations are based on the networks of connections perceived between things rather than inherent qualities of the things themselves.

Individual items are nodes in these collection networks. Museums Victoria is responsible for close to seventeen million artifacts, specimens, photographs and documents collected over 160 years, with these currently managed by departments operating under the broad umbrellas of Humanities and Natural Sciences. Some items exist as part of multiple collections. Perhaps the most prom-

https://doi.org/10.1515/9783110337419-008

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

inent example at Museums Victoria is Phar Lap, Australia's most famous racehorse, a biological specimen, a cultural, historical and sporting icon, and the centerpiece of the Phar Lap Collection (Churchward 2010).

But, despite significant advances over the past fifty years, collection documentation in museums often remains focused less on networks supporting diverse requirements and more on discrete items, hierarchical structures and clear boundaries between disciplines. Technology is a contributing factor. Where Satterthwait believes that collection formation is "a sociocultural process" (Satterthwait 2008, 30), an examination of the tools and practices used by collecting institutions shows that the process is also a socio-technical one.

In the First Peoples section of Melbourne Museum is a large gallery containing more than 600 historic and contemporary pieces showing the diversity of Aboriginal and Torres Strait Islander cultures in Australia (Harford 2013). Coiled at the bottom left corner of the Working Country case is an item called *kopoi*, or 'dugong rope', plaited light brown Hibiscus fiber attached to a pair of straight wooden harpoon heads. One of the harpoons is marked with a handwritten black identifier, DT 3324, referring to Donald Finlay Fergusson Thomson (1901–1970) who collected the item on his first expedition to Cape York in 1928–1929.

When the Donald Thomson Collection was transferred to the National Museum of Victoria (predecessor to Museums Victoria) in 1973, the organization claimed it as "probably the most important [collection] to come under the control of a public institution for many decades" (National Museum of Victoria Council 1973). Although often discussed as a singular entity, the Thomson Collection consists of many potential groupings depending on perspective. With regard to ownership, it includes some 7,200 artifacts and 2,500 natural history specimens on long-term loan from the University of Melbourne, and around 11,000 photographic images, as well as audio visual material, field notes, diaries, transcriptions and drawings (Thomson's 'literary estate') which remain the property of the Thomson family. Both sets of material were transferred to the National Museum as part of a three-way agreement between the Museum, the University and Thomson's widow Mrs. Dorita Thomson.

The collection includes items related to (1) anthropology, (2) linguistics, (3) genealogy, (4) botany, (5) zoology, (6) ornithology and (7) ecology. Geographically, Thomson collected material from Cape York Peninsula, Arnhem Land, Central Australia, the Solomon Islands and West Papua over his forty-year career, from 1928–1968. It includes three-dimensional museum objects, unpublished archival material, and publications. Curator Lindy Allen sees it as "a series of field collections from different periods and places and often collected with somewhat different rationales" (Allen 2008).

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

The richness of the collection has resulted in its use across a range of disciplines. Early in his career Thomson himself published on both zoology and anthropology (Thomson and Peterson 1983), though as his career progressed he became more associated with the latter. Due to concerns about the safety of the collection few others were able to access the material during his lifetime (Morphy 2015). Following Thomson's death in 1970 there have been exhibitions dedicated to Thomson photographs, bark paintings and artifacts, and the dugong rope is one of more than seventy artifacts from the collection currently on display at Melbourne Museum.

In addition to continuing museum-based research, collection material has been used as part of Indigenous land claims (Neate 2001), as source material for work on Aboriginal architecture (Memmott and Fantin 2007) and studies of Indigenous companion animals (Philip and Garden 2016), and as inspiration for the film *Ten Canoes* (Hamby 2007; Rutherford 2011), among other projects. Members of Aboriginal communities, particularly from Arnhem Land and Cape York, have also been regular visitors to Museums Victoria, engaging with artifacts, photographs and descriptive content from what some call 'Thomson Times' (Hamby 2007; Allen and Hamby 2011; Rutherford 2011). When recognized by UNESCO as part of the Australian Memory of the World register in February 2008, the entry noted that for some communities the collection was "the only record of their heritage" (Australian Memory of the World Program 2008, par. 5).

The strength of the collection in supporting such work is not just in its individual elements, but also in the aggregate. Some see a distinction between artifacts and documentary context, viewing the former as collection items and the latter as related contextual material. Satterthwait (2008) is one, writing of anthropological collections:

> A distinction can be made between information that might be obtained *about* a collection and what is absent from it, and information that might be obtained *from* a collection. The first corresponds to information obtainable from sources external to the collection itself, such as published accounts, archival documents, photographs ... the collector's notes and diaries, and, most important of all, the people who created and used the objects in the first place. (Satterthwait 2008, 32)

Many museums organize their holdings in this way, with archival material and library collections external to 'the museum collection', managed and documented separately (Jones 2015).

Museums Victoria is no different. The State Heritage Collection is conceptually and physically separate from the organization's historical archives and current records, and the two are managed in different databases with few links be-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

tween them (McNulty 2000; Borg 2016). The Museum's library is separate again. However, at Museums Victoria such divisions are often based as much on provenance as format. Had Thomson been a staff member of the National Museum of Victoria his archival documents, photographs, diaries and other material may well have gone into the organizational archives, with his artifacts and specimens deposited in the State Heritage Collection.

As an external collector, material was seen as part of the Donald Thomson Collection regardless of format, and Satterthwait's (2008) distinction between material *about* the collection and material *in* it starts to blur. The description held by the Museum Metadata Exchange highlights this point, noting that "much of the value of the Collection lies in its careful interweaving of various sources of information and documentation and images" (Museum Victoria – The Donald Thomson Collection 2012). Making such interweaving visible to users requires detailed and effective documentation.

Dr. Ray Marginson was chairman of the Donald Thomson Collection Administration Committee for forty years, from 1973 when the collection was transferred to the Museum until 2013. The loan to the National Museum of Victoria took place in large part because the University of Melbourne did not have the space or resources to curate, sort and preserve the collection. Cataloging, housing and researching the collection was, he later recalled, a "herculean task" (Ray Marginson, quoted in Inglis 2013, 36). But rather than a singular, finite event, documenting Thomson has involved a sequence of interlocking phases and technologies, and is a process which continues today.

Prior to his death, the collection and its documentation remained under the control of Donald Thomson himself. In 2008, Museum Victoria CEO Patrick Greene commented that Thomson set the standard for anthropologists in his methodical record-keeping and documentation (Museum Victoria 2008). Lindy Allen has written in more detail of his practices, which included tags attached to his ethnographic material as well as his biological specimens. These identified the item and provided details such as its indigenous name, the group from which the item was collected or with which it was associated, information on the place and date of collection, and sometimes additional context (Allen 2008). Thomson also kept journals and detailed field notes, often accompanied by sketches, and captured the preparation of materials and the manufacture of some objects with series of photographs (Allen and Wrench 2005).

When the collection came to the Museum in 1973 these layers of content were not readily discoverable or accessible. Fortuitously, Judith Wiseman came too. She had been Thomson's secretary, technical assistant, and finally his administrative assistant during the last years of his life. As a result, Wiseman held significant implicit knowledge about the collection and quickly proved to be in-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

dispensable (National Museum of Victoria Council 1973; National Museum of Victoria Council 1974), not least because, as Christine Hogarth noted, "in many instances, Wiseman is the only one who can decipher Thomson's handwriting" (Hogarth 1980, 6).

With work lead by Wiseman, Thomson's tags and other documentation became the primary source for cataloging efforts. Documents from the time note that Mammalian, Herpetological and Ornithological sections of the collection were transferred to the relevant Museum departments, with the remainder of the collection stored sorted by regions and types (National Museum of Victoria Council 1973). The intention was to produce duplicate typed catalog cards for each artifact to facilitate access to data and specimens (Wiseman 1973).

By October 1975, 1,500 cards had been completed from an estimated 5,750 required. Fifteen months later Nicolas Peterson, writing to the University of Melbourne's Deputy Vice-Chancellor Professor David Caro, already believed the collection to be very well documented thereby providing the "rare opportunity" to work with a variety of material culture artifacts which were associated with behavioral data (Peterson 1977). But Wiseman estimated effective documentation of the whole collection would take until at least 1980.

Capturing the interweaving of the various objects, specimens, images and text in the Thomson Collection required more than just discrete item records for artifacts. Wiseman also worked on deciphering and transcribing thousands of pages of Thomson's field notes into typed foolscap pages, a task that took until 1988 to complete (Annual Report 1987–1988). She and her colleagues were also printing and labeling photographs (Inglis 2013), and working on cross-referencing artifacts, field notes, photographs and other material (Menkhorst 1977; Wiseman 1978).

Early work on the collection was necessarily manual, analog and labor intensive; but museum technology was starting to change. Apart from the Thomson transfer, 1973 marked another significant event in the history of anthropological collections in Australia. The Department of Anthropology at Sydney's Australian Museum were concerned about their ability to retrieve data related to their collection (Lampert and Specht 1979), so they looked to the emerging use of computers in American and European museums (Bergengren 1971; Lewis 1971; Jones-Garmil 1995) for the solution. Over the next four years, at a cost of ten dollars per record, the Australian Museum entered 3,200 records in the Australian Ethnographic File (Lampert and Specht 1979), the first example of computer-assisted museum cataloging in the country.

The Australian Institute of Aboriginal Studies (AIAS) quickly saw the potential of this approach. Shortly after the Australian Museum started their ethnographic file, AIAS, which funded the first years of work on the Thomson Collec-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

tion, were suggesting catalogers use standard forms and codes, to achieve uniformity across the whole of Australia (West 1978). On May 2, 1977 Jan Menkhorst started as a Graduate Cataloger with the Donald Thomson Collection. In line with AIAS recommendations she prepared computer sheets, 1,050 for Arnhem Land artifacts in her first four months (Menkhorst 1977).

But despite the terminology, the process was still far from digital. The sheets were handwritten, with some fields (such as "Functional qualifier" and "Makers' category") using numeric codes rather than text. Once completed, a typist prepared catalog cards that also featured these codes. The idea was that when the collection data was aggregated with content from the Australian Museum and elsewhere these standardized fields would allow the grouping of, for example, weapons using Functional qualifier thirty-two: Fighting and hostility.

Seven years after the collection arrived at the Museum, all the artifacts had finally been sorted and cataloged, 7,000 of the 10,000 photographs copied, and 4,000 of 7,000 sheets of field notes copied, deciphered and typed (Hogarth 1980). Unfortunately, the Australian Museum project had stalled due to lack of funds, and AIAS grants for similar work terminated (Marginson 1978; Donald Thomson Collection Administration Committee 1978). With the collection documentation now consisting of several thousand physical cards, thoughts turned to a more effective way of disseminating information about the collection to researchers.

With desktop computing still several years away, the Thomson Collection Administration Committee settled on microfiche. But imaging the existing catalog cards was far from ideal because of the data that had been coded. While these numeric identifiers could facilitate computer processing, if simply reproduced they would be difficult for researchers to read and interpret (McCahon 1980). Therefore, as the coding system used was in accordance with the Australian Museum, the decision was taken to use their system with minor adaptations. What followed was, by today's standards, a convoluted process that demonstrates the practical complexities of early museum computerization projects (1) keying the required coding sheets, (2) sending them to Singapore for transfer to magnetic tape, (3) utilizing the Commonwealth Scientific and Industrial Research Organization's mainframe computing services (CSIRONET) to produce and print the catalog text, and (4) finally, manually affixing photographs items to these printed sheets. Only then could the microfiche itself be produced. Work on transferring collection data onto data entry sheets commenced in March 1981 (National Museum of Victoria 1981); the whole process took six years.

Aboriginal Artefacts in the Donald Thomson Collection – A Microfiche Catalogue contains a thirty-four page printed booklet and ten microfiche sheets. As indicated by the title it does not describe specimens, field notes, photographs

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

or other material, but within the limitations of its scope and technology the catalog has great strengths. An index allows for searches by function and category of artifact, geographic location, or tribal group. (This information is not available for all items. The dugong rope, for example, cannot be found by location, though it is now known to be from Lockhart River). Items are also sorted by geographic area (1) Cape York Peninsula, (2) Arnhem Land, and (3) Desert and then by function.

For example, the dugong rope is on the first microfiche sheet from Cape York Peninsula, grouped with other harpoon equipment under the broader category of "Hunting and Fighting," which is part of the top-level category "Utensils and Implements." Its catalog entry reads:

> DT3324
> **Harpoon heads and rope**
> Two single pointed wooden harpoon heads with wooden barbs similar to those in Plate Y7 but attached to a length of two ply rope. Used for dugong hunting.
> Cape York Peninsula. 1928–1933.
> Refs: TPH 2813–28; TPUB 5.

'Plate Y7' is an image of the following:

> DT3328
> **Harpoon head**, wood with metal barb
> A single pointed wooden head with metal barb. Barb is attached with vegetable fibre and a small piece of wax/resin has been attached near the proximal end of the head. Used for hunting dugong.
> L. 319mm
> Cape York Peninsula. 1929. Gan-ganda.
> Refs: TFN 192; DT BOT 381A; TPUB 5.

As seen here, description in the microfiche catalog is focused on materials, location and basic function. If the user wants to explore further, the explicit cross-references then provide a pathway to additional resources. TPH refers to photographs. For example, TPH 2813 is an image of a Kuuku-ya'u man from the Cape York Peninsula preparing plant material for making rope, and TFN 192 is a set of field notes related to the harpoon head described.

TPUB 5, related to both items, refers to one of Donald Thomson's publications listed in Appendix 1 of the printed booklet accompanying the microfiche guide. The article, "The Dugong Hunters of Cape York" (1934), includes information on rope materials and manufacture, and an evocative description of a hunt. Thomson writes of the harpooners "leaping bodily overboard" to drive the harpoon into the body of the dugong before throwing themselves out of the path of the rope, which uncoils as the animal takes off. After towing the boat

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

for some distance, the dugong tires and comes up for breath and is finally subdued (Thomson 1934, 242–247, 255–257).

These relationships between elements were seen as central to the collection's value. As Nicolas Peterson wrote in 1983:

> Its importance lies not only in its size and comprehensiveness ...but in the superb documentation of the items and the interrelationships between objects, photographs and notes which make it unique among collections from this period and a rich resource both for anthropologists and archaeologists, and as an historical archive for the people of Cape York and Arnhem Land. (Thomson and Peterson 1983, 13–14)

But despite its strengths, and the use of computer processing in its publication, the microfiche catalog also retained the limitations of print. Indexes and sort orders remained static and limited to particular subjects, with these biased toward culture and anthropological enquiry. A user interested in hunting could find everything related to their topic relatively quickly, but another interested in dugongs would potentially need to skim the entire guide, or make educated guesses about practices and subjects that might include reference to the animal. There was also no way of incorporating the new relationships and other information uncovered by communities, curators and researchers who continued to work with the prominent collection. The guide was a static snapshot, whereas Judith Wiseman's report "Work still to be done – March, 1978" concluded that the cross referencing of object could continue on forever (Wiseman 1978).

In 1987, the same year the microfiche catalog was completed, the first database for Aboriginal collections was established at the museum in the midst of a period of significant organizational and technological change. The Science Museum of Victoria had started its own microfiche catalog in 1976 (Hart and Hallett 2011) and the National Museum of Victoria began collaborating with the University of Melbourne on potential new systems in the early 1980s. When the two institutions merged in 1983 to form the Museum of Victoria (later just Museum Victoria, and now Museums Victoria) these technology streams came together. Shortly thereafter, Titan was introduced by Knowledge Engineering Pty. Ltd., a company founded to continue the commercial development of projects undertaken by the University's Computer Science Department. Titan led to Texpress, Knowledge Engineering became KE Software, and in 1997 Museum Victoria invested in a new concept pitched by the company, KE EMu. They were one of the first to do so, and would become instrumental in the system's development.

By this time, the Indigenous Cultures department had four databases (1) Ethno, which contained records about artifacts, (2) Audio Visual, (3) Manuscript, and (4) Archaeology (which was limited). The 'ethno' database alone held more than 90,000 records, while across the museum there were more than

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

two hundred Titan and Texpress databases, around fifty of them considered high priority for transfer to the new EMu system. At institutional level it was a process that would take around ten years.

Looking at KE EMu today, some of the description is comparable to that seen in the microfiche catalog. Returning to the dugong rope, its record has the same registration number (DT 003324), is registered as part of the Donald Thomson Collection, and includes a brief physical description of the object. Other elements are new. The large number of fields and tabs in KE EMu allows space for information about the loan from the University of Melbourne, details of updates to language names, a barcode reference to assist with location tracking, and more. Where the microfiche only referred to a single black-and-white image of a similar object, KE EMu contains ten color photographs accessible via the Multimedia tab.

Cross-references have also been expanded, though their roots in earlier work are still visible. The 'Notes' field starts with: "Reference/Publications: TPUB 5 Thomson, D.F. 'The Dugong Hunters of Cape York' Jour.R.A.A. p. 250. vol 64. 1943." While the TPUB 5 reference has been retained, the user no longer has to go elsewhere to look up the code. And 'Secondary Comments' reads: "TPH 2813–2828 – Related objects: 3315–3317,3321,3323." The photographic references are repeated from the earlier catalog, but the object references have since been added. None of these cross-references are available as active links. Though search functionality makes it a relatively straightforward task, users still have to manually look up related objects, photographs and other material.

There is another aspect of the current system worth considering. Whereas the microfiche catalog was a publication, making it accessible to many people in Australia and internationally, KE EMu remains an in-house system. External researchers, communities and the general public can only access information about the Thomson Collection by physically visiting the Museum's exhibition spaces, or searching for material online.

The Many Nations gallery space, mentioned earlier in this chapter, contains innovative digital displays with touch screens that visitors can use to find out more about individual objects. Here, the dugong rope is presented with fielded information on its (1) date, (2) maker, (3) clan/language group, (4) location (including a map), (5) materials, and (6) the item's registration number. This is followed by narrative text providing information on dugong hunting, rope manufacture, and the connection between this and the practices of the Wulmpano (Totemic Ancestors). Much of this looks to be based on Thomson's 1934 article, though no reference is provided. One of Thomson's photographs is also included, showing a man finishing the end of a rope ready to attach to a harpoon, but no reference number is visible. Therefore, while the content populating the

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

display looks to have been developed utilizing the cross-references to Thomson's publications and photographs, the links are no longer explicit. Those who wish to find out more will need to do their own research.

Online, users seeking information about the dugong rope have a different experience again. Museums Victoria Collections provides eight high-quality color photographs, key metadata about the artifact, a physical description, and a short summary about harpooning that is more concise and formal than either the in-gallery display or Thomson's own published work. Beneath this is a list of fielded data about collections and collectors, dates, locations, classifications and more; however, the references to related photographs and publications are not included, nor are the references to other objects. Similarly, looking at DT 3328, the harpoon head for which the microfiche catalog details were quoted above, there is no reference to field notes or botanical specimens.

Aside from its reputation as one of the most comprehensive and significant collections of Aboriginal cultural heritage material in the world (Museum Victoria 2013), the Thomson Collection is also a rarity in terms of documentation. When asked which collection at Museums Victoria could be considered an exemplar in this area, Senior Curator Lindy Allen nominated Thomson due in large part to the work of Judith Wiseman on transcribing, documenting and cross-referencing material. When asked if there were any other collections with relationships across materials described in this way, Allen stated: "I've not come across one. Not here or anywhere else" (Allen 2016, personal interview).

In the case of Thomson, the resulting ability to locate documentary material and photographs is particularly pertinent because he published comparatively little (Rigsby and Peterson 2005; Allen 2016), making this material central to understanding artifacts. A 1975 funding proposal prepared by the Museum stated that it was essential that copies of the material, including photographs owned by Mrs. Thomson, remain with the collection to enable complete interpretation of the total collection (National Museum of Victoria 1975). What remains unsaid, but is equally essential, is that these materials remain discoverable.

Evidence suggests past practice was based on this aim. Substantial resources were dedicated to documenting the Thomson Collection from 1973 to 1980, followed by additional resources from 1981–1987 preparing and publishing the microfiche catalog. Few collections have been the focus of such a sustained effort; but the fact that this relational data is rarely available means that the design of online collections sites has moved in a different direction.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

Conclusion

Museums Victoria Collections, as with comparable sites around the world, is centered on search. Users can quickly find material both within the Thomson Collection and across collections, and can then filter based on geographic location, item type, and other facets. Finding diverse collection items related to dugongs, which in the past would have involved significant work, now takes seconds. But discovering explicit connections (two objects collected at the same time and described on a single page of field notes, the set of material described in a published article, or a photograph of an artifact being manufactured) is not only left up to the user, it is a process that then needs to be repeated by each subsequent user interested in the same item.

Many discussions about museums in the digital age include mention of decentralization, distribution of content, and the challenge and potential of networks (Gere 1997; Cameron and Robinson 2007; Byrne et al. 2011; Harris 2013). Though the Thomson Collection's documentation contains much of the data required to support such an approach, the systems used to capture, manage and disseminate that information to the public limit the utility and accessibility of that data.

Even with Thomson there is more that could be done to expand the relationship data already captured. The cross-references could be converted to active links between items, and additional data could be added to indicate why the reference exists. For example, a photograph showing the actual manufacture of a particular object could include a reference to this fact as part of the cross-reference, to distinguish it from a relationship to a photograph of a similar but different object. Similarly, a cross reference to a publication could be because that publication includes detail about a specific artifact, or because it includes narrative material that helps the user understand the cultural context and use of the type of item they are examining. The limits to this work would be more related to resources than technology. As Wiseman noted, identifying and enriching cross-references could go on forever.

But to realize the benefits of this work we need to reconceptualize the way we capture and disseminate collection description more broadly. Text search and keywords have opened up new possibilities, but this does not mean the relationships found in the Thomson microfiche catalog no longer have value. Contemporary digital practice also needs to consider collection networks and the webs of meaning between items, including archives and published works, providing richer information to users and more effectively sharing the knowledge held by institutions with the broader public.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

References

Allen, Lindy. 2008. "Tons and Tons of Valuable Material: The Donald Thomson Collection." In *The Makers and Making of Indigenous Australian Museum Collections*, edited by Nicolas Peterson, Lindy Allen, and Louise Hamby, 387–418. Carlton, Vic: Melbourne University Publishing.

Allen, Lindy. 2016. Museum Victoria – Lindy Allen Interviewed by Mike Jones.

Allen, Lindy, and Louise Hamby. 2011. "Pathways to Knowledge: Research, Agency and Power Relations in the Context of Collaborations between Museums and Source Communities." In *Unpacking the Collection: Networks of Material and Social Agency in the Museum*, edited by Sarah Byrne, Anna Clarke, Rodney Harrison, and Robin Torrence, 209–29. One World Archaeology. New York: Springer. http://dx.doi.org.ezp.lib.unimelb.edu.au/10.1007/978–1–4419–8222–3.

Allen, Lindy, and Rosemary Wrench. 2005. "Donald Thomson Collection Development Plan." Museum Victoria. CM/035/30~23.

Australian Memory of the World Program. 2008. "Donald Thomson Ethnohistory Collection." *UNESCO National Committee of Australia – Memory of the World*. http://www.amw.org.au/register/listings/donald-thomson-ethnohistory-collection.

Bergengren, Göran. 1971. "Automatic Data Processing in the Registration of Museum Collections in Sweden." *Museum International* 23 (1): 53–58. doi:10.1111/j.1468–0033.1971.tb01751.x.

Borg, Belinda. 2016. Museum Victoria – Belinda Borg Interview by Michael Jones.

Byrne, Sarah, Anna Clarke, Rodney Harrison, and Robin Torrence, eds. 2011. *Unpacking the Collection: Networks of Material and Social Agency in the Museum*. One World Archaeology. New York: Springer. http://dx.doi.org.ezp.lib.unimelb.edu.au/10.1007/978–1–4419–8222–3.

Cameron, Fiona, and Helena Robinson. 2007. "Digital Knowledgescapes: Cultural, Theoretical, Practical, and Usage Issues Facing Museum Collection Databases in a Digital Epoch." In *Theorizing Digital Cultural Heritage: A Critical Discourse*, edited by Fiona Cameron and Sarah Kenderdine, 165–91. Cambridge, Mass: MIT Press.

Churchward, Matthew. 2010. "Phar Lap Collection." *Museum Victoria Collections*. http://collections.museumvictoria.com.au/articles/3205.

Donald Thomson Collection Administration Committee. 1978. "University of Melbourne – Donald Thomson Collection Administration Committee – Meeting No. 2 – 1978." National Museum of Victoria File, Administration – General – The Donald Thomson Collection, Administration Committee – Minutes of Meetings 1973–1982. Museum Victoria Archives.

Gere, Charles. 1997. "Museums, Contact Zones and the Internet." In *Museum Interactive Multimedia 1997: Cultural Heritage Systems Design and Interfaces. Selected Papers from ICHIM 97*, 59–66. Le Louvre, Paris, France: Archives & Museum Informatics. http://www.archimuse.com/publishing/ichim97/gere.pdf.

Hamby, Louise. 2007. "Thomson Times and Ten Canoes (de Heer and Djigirr, 2006)." *Studies in Australasian Cinema* 1 (2): 127–46. doi:10.1386/sac.1.2.127_1.

Harford, Sonia. 2013. "First Peoples Share Rich Cultural Legacy." *The Age*, September 6. http://www.theage.com.au/entertainment/art-and-design/first-peoples-share-rich-cultural-legacy-20130905–2t7ub.html.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

Harris, Clare. 2013. "Digital Dilemmas: The Ethnographic Museum as Distributive Institution." *Journal of the Anthropological Society of Oxford Online* V (2): 125–36.

Hart, Tim, and Martin Hallett. 2011. "Australian Museums and the Technology Revolution." In *Understanding Museums: Australian Museums and Museology*, edited by Des Griffin and Leon Paroissien. National Museum of Australia. http://nma.gov.au/research/understanding-museums/THart_MHallett_2011.html.

Hogarth, Christine. 1980. "Donald Thomson Collection – Completion of Basic Curation." National Museum of Victoria File, Administration – General – The Donald Thomson Collection, Administration Committee – Minutes of Meetings 1973–1982. Museum Victoria Archives.

Inglis, Alison. 2013. "Retirement and Recollection: Dr. Ray Marginson AM and the Donald Thomson Collection." *University of Melbourne Collections*, no. 13 (December): 35–38.

Jones-Garmil, Katherine. 1995. "Museums in the Information Age." In *Hands On: Hypermedia & Interactivity in Museums: Selected Papers from the Third International Conference on Hypermedia and Interactivity in Museums (ICHIM 95 / MCN 95)*, edited by David Bearman, 2:1–12. Archives & Museum Informatics. http://www.archimuse.com/publishing/ichim95_vol2/jones-garmil.pdf.

Jones, Michael. 2015. "Artefacts and Archives: Considering Cross-Collection Contextual Information Networks in Museums." In *Museums & the Web: Selected Papers and Proceedings from Two International Conferences*, edited by Nancy Proctor and Rich Cherry, 123–35. Melbourne, Australia: Museums and the Web. http://mwa2015.museumsandtheweb.com/paper/artefacts-and-archives-considering-cross-collection-knowledge-networks-in-museums/.

Lampert, R. J., and Jim R. Specht. 1979. "Computer Cataloging of Anthropological Collections." *Kalori*, no. 56 (June): 39–44.

Lewis, Geoffrey. 1971. "An Interdisciplinary Communication Format for Museums in the United Kingdom." *Museum International* 23 (1): 22–26. doi:10.1111/j.1468–0033.1971.tb01748.x.

Marginson, R. D. 1978. "Computerization of Ethnographic Collections – Thomson." National Museum of Victoria File, Administration – General – The Donald Thomson Collection, Administration Committee 1975–1982. Museum Victoria Archives.

McCahon, J. 1980. "Progress Report – Donald Thomson Collection Microfilm Catalogue." National Museum of Victoria File, Administration – General – The Donald Thomson Collection, Administration Committee – Minutes of Meetings 1973–1982. Museum Victoria Archives.

McNulty, Alison. 2000. "Museum Victoria Information and Records Management." *Archives and Manuscripts* 28 (2): 152–54.

Memmott, Paul, and Shaneen Fantin. 2007. "Symbolism and Meaning in Aboriginal Architecture." In *Gunyah, Goondie + Wurley: The Aboriginal Architecture of Australia*, 232–256. St Lucia, Qld: University of Queensland Press.

Menkhorst, Jan. 1977. "Report No. 1 – Report to: The Principal, Australian Institute of Aboriginal Studies." National Museum of Victoria File, Administration – General – The Donald Thomson Collection, applications and receipt of grants 1972–1979. Museum Victoria Archives.

Morphy, Howard. 2015. "Thomson, Donald Finlay Fergusson (1901–1970)." In *Australian Dictionary of Biography [online Edition]*. Canberra: National Centre of Biography, Australian National University. Accessed July 31. http://adb.anu.edu.au/biography/thomson-donald-finlay-fergusson-11851.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

Museum of Victoria. 1987. "Museum of Victoria – Annual Report 1986/7." Museum of Victoria.
Museum Victoria. 2008. "Donald Thomson Ethnohistory Collection recognized on UNESCO list." Media Release.
Museum Victoria. 2013. "Donald Thomson: Museum Victoria." *Museum Victoria*. December. https://museumvictoria.com.au/collections-research/our-collections/indigenous-cultures/donald-thomson/.
"Museum Victoria – The Donald Thomson Collection." 2012. *Museum Metadata Exchange*. May 30. http://museumex.org/oai/mv/2785.
National Museum of Victoria Council. 1973. "Application to the Australian Institute for Aboriginal Studies for Funds to process the Donald Thomson Collection." National Museum of Victoria File, Administration – General – The Donald Thomson Collection, applications and receipt of grants 1972–1979. Museum Victoria Archives.
National Museum of Victoria Council. 1974. "Application for a Grant to the Australian Institute for Aboriginal Studies." National Museum of Victoria File, Administration – General – The Donald Thomson Collection, applications and receipt of grants 1972–1979. Museum Victoria Archives.
National Museum of Victoria. 1975. "Application for a Grant from The William Buckland Foundation." National Museum of Victoria File, Administration – General – The Donald Thomson Collection, applications and receipt of grants 1972–1979. Museum Victoria Archives.
National Museum of Victoria. 1981. "Annual Report / National Museum of Victoria." National Museum of Victoria.
Neate, Graeme. 2001. "Donald Thomson's Legacy and the Resolution of Indigenous Land Claims in Australia – Opening Address." In The University of Melbourne. http://www.nntt.gov.au/Information%20Publications/Speeches%20Donald%20Thomson%20legacy%20Neate%20July%202001.pdf.
Peterson, Nicolas. 1977. Letter to Professor Caro, 14 January 1977. National Museum of Victoria File, Administration – General – The Donald Thomson Collection, Administration Committee 1975–1982. Museum Victoria Archives.
Philip, Justine, and Don Garden. 2016. "Walking the Thylacine: Records of Indigenous Companion Animals in Australian Narrative and Photographic History." *Society & Animals* 24 (1): 34–62. doi:10.1163/15685306-12341386.
Ramsay, E. G. 1987. *Aboriginal Artefacts in the Donald Thomson Collection a Microfiche Catalogue*. Melbourne: University of Melbourne and the Museum of Victoria.
Rigsby, B., and Nicholas Peterson. 2005. *Donald Thomson: The Man and Scholar*. Canberra: Academy of the Social Sciences in Australia with support from Museum Victoria.
Rutherford, Anne. 2011. "Ten Canoes and the Ethnographic Photographs of Donald Thomson: 'Animate Thought' and 'the Light of the World.'" *Cultural Studies Review* 18 (1): 107–37.
Satterthwait, Leonn. 2008. "Collections as Artefacts: The Making and Thinking of Anthropological Museum Collections." In *The Makers and Making of Indigenous Australian Museum Collections*, edited by Nicolas Peterson, Lindy Allen, and Louise Hamby, 29–60. Carlton, Vic: Melbourne University Publishing.
Thomson, Donald F. 1934. "The Dugong Hunters of Cape York." *The Journal of the Royal Anthropological Institute of Great Britain and Ireland* 64: 237–64. doi:10.2307/2843809.
Thomson, Donald F., and Nicolas Peterson. 1983. *Donald Thomson in Arnhem Land*. Melbourne: Currey O'Neil.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

West, Alan L. 1978. "Summary statement on funding of salaries – The Donald Thomson Collection." National Museum of Victoria File, Administration – General – The Donald Thomson Collection, Administration Committee 1975–1982. Museum Victoria Archives.

Wiseman, Judith. 1973. "Report to: The Vice-Chancellor." National Museum of Victoria File, Administration – General – The Donald Thomson Collection, applications and receipt of grants 1972–1979. Museum Victoria Archives.

Wiseman, Judith. 1978. "The Donald Thomson Collection – Work still to be done – March, 1978." National Museum of Victoria File, Administration – General – The Donald Thomson Collection, Administration Committee 1975–1982. Museum Victoria Archives.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

Jennie Choi and Giovanna Fiorino-Iannace

8 Information Management Systems at the Metropolitan Museum of Art

Introduction

Socio-technical interactions occur when people, information and technology interconnect, and can be studied to analyze the changes in technology that have transformed and enhanced the experience of individuals visiting libraries, archives, and museums (LAMs). Given that there is a continuous demand for comprehensive access to information resources spanning all types of collecting agencies, it is important to understand that cultural institutions, such as LAMs, are unique repositories of information. They are fundamentally different institutions, and require wide-ranging approaches to information organization and access behind the scenes. Likewise, it is essential to remember that while LAMs are entrusted with preserving and protecting their artifacts, they are also expected to incorporate advanced technology toward more effective preservation and more efficient accessibility.

It is instructive for researchers to understand how data is captured and processed in what appears to be an immediate response to an inquiry, as such a response actually involves a great amount of time on the part of the information professional. Museum information staff in particular need to have a strong knowledge of cataloging and classification systems for art objects, and must be capable of transferring written descriptive information to image databases and content management systems for expedient retrieval by students and researchers.

The Metropolitan Museum of Art (The Met) may be looked upon as an example. Located in New York City, The Met is the largest art museum in the United States and consists of seventeen curatorial departments. Its collection spans over 5,000 years and represents all areas of the world. As part of a visitor outreach feature, The Met uses technology that requires the collaboration of several departments. The manner in which LAMs in general approach technology varies for many reasons, ranging from budgetary limitations that challenge institutional priorities, to the availability of trained staff who can maintain the work flow. Nonetheless, it is impossible for institutions to avoid the onset of and full engagement in the digital age. The goal of providing accurate descriptive information about The Met's extensive collection of art is an ongoing process that

https://doi.org/10.1515/9783110337419-009

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

has evolved from using typed catalog cards at the beginning of the twentieth century to the development of computer-based platforms in the mid-1990s that were capable of introducing this data to a wider audience.

Today The Met is comprised of three branches (Fifth Avenue, The Cloisters and Breuer). The Fifth Avenue location is considered the museum's main location, while The Met – Breuer, on the corner of 75th Street and Madison, features the museum's modern and contemporary art collection, and The Met – Cloisters, situated in Fort Tryon Park, maintains the museum's Medieval Art holdings. The museum is three museums in one and maintains one central collections management system. The system combines all data relating to the basic description of every object in its collection, as well as a representative image of each object, and information on historical relevance, current condition, and storage location. The descriptive elements are accessible via the Met's website, which is driven by a robust image database system monitored by a team of collections information specialists. The Met's use of this comprehensive system began almost a quarter century ago with the development of the museum's Antonio Ratti Textile Center.

Selecting an image database for the Antonio Ratti Textile Center

The necessary process of selecting the appropriate long-term collections management system can be labor-intensive. It involves a critical evaluation of how the institution uses, manages and assigns data to the objects it owns. This chapter discusses the evolution of the use of technology with the introduction of an image database at The Met, which began over twenty years ago with the unification of the textile collection in one resource center. This turning point in The Met's history changed the way catalog information was accessed by the staff as well as the way it was presented to and accessed by the public audience via the museum's website (launched in 1995).

The first portion of the chapter will explore how the creation of the Antonio Ratti Textile Center, The Met's main repository for the storage and study of its historical textile collection, paved the way for the eventual implementation of one consolidated database for the entire art collection. The steps leading to one central collections management system were often challenging, even with the support of the department of Information Technology and Systems, and required careful planning on the part of the Collections Information Management

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

team entrusted with this task. It is important to keep in mind that in many instances museum staff members did not have access to a computer, and in the early 1990s there were no staff email accounts. What may seem to researchers today like a standard collections management endeavor was a revolutionary moment for The Met.

In 1994, as part of the preparation for the opening of the textile study and storage facility in the Fifth Avenue branch to be known as the Antonio Ratti Textile Center (RTC or the Center), The Met was faced with the challenge of uniting the records for thirteen curatorial departments into one central database system devoted to the museum's collection of textiles, tapestries, and carpets (excluding the costumes and accessories belonging to the Costume Institute). The level of basic collections data prior to the use of a collections management system varied across the curatorial departments, some of which were documenting the collection on Microsoft Word and some of which had yet to convert information from paper to an electronic system.

The Ratti Center was in a unique position, directly involved with the development of the database's architecture and specific features relating to the tracking of objects, and also with the descriptive elements of the system. As part of the initial team of staff members working closely with Jay Hoffman, President and CEO of Gallery Systems, creator of the Windows-based version of The Museum System (TMS), the RTC staff was able to provide constructive feedback about data entry and the retrieval of information in TMS during the development of the Center.

The Museum's retail shop was using a DOS-based version of TMS at the time, as was the department of Twentieth Century Art (now known as the Modern and Contemporary Art Department). It was not until the planning for the Ratti Center that the museum started considering the use of one system to unite descriptive, visual and location data. The first collections management system that was considered was from the Argus Interactive Agency, but this system was not chosen because it did not include the option to add images to the data records.

Gallery Systems was on the rise as a company that offered a collection management system that also permitted the incorporation of multiple images and videos. Fortunately, Jay Hoffman was based in the Ratti Center while he and his team designed the new Windows-based version of TMS. Before the option to create customizable screens became available, the Center made recommendations for the layout of the main data entry screen and shared preferences for report designs. Several staff members were trained to use Crystal Reports, which was a way of creating customized reports from the different data tables that existed (which were not many in 1994). The Center alone manually entered close to

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

40,000 records in TMS, including objects stored in the study and storage facility itself, and textiles, carpets and tapestries physically located in various departmental storerooms throughout the museum.

The standardization of terms and the way the many curatorial departments used data entry fields, such as "Culture" and "Period" may be considered the most challenging part of the process as the museum attempted to enter the basic catalog information as it was shown on the actual catalog cards. Fields such as object name, classification (representing the technique), culture, country, and date, were among the most debated, and required much discussion because each curatorial department had a different use for these fields. The recommended use of these fields did not always conform to the way information was being presented in the departments. This non-conformity needed to be addressed and the TMS advisory group, consisting of Information Systems staff as well as curators and conservators, assisted in adjusting the data in order to help standardize the use of each field. The goal was to provide data on the collection in a consistent format for ease of use of the database and retrieval of information.

In 1996, a year after the official opening of the Ratti Center, The Met chose to adopt TMS for all its content management needs, and each of the seventeen curatorial departments began using its own separate version of the system. It was because of this museum-wide use that in 2013, when the data from all curatorial departments was consolidated into one database, most of the information was already in the system. TMS at The Met today contains the entire range of media types, including paintings, sculpture, decorative arts, drawings, prints, musical instruments, arms and armor in one all-encompassing collections management system.

To date, there are almost half-a-million records in the museum's TMS database, with some departments owning under 3,000 art works and others responsible for more than 129,000 objects. The maintenance of all these records is carefully monitored by each individual department's collections management staff, who are entrusted with ensuring the safety and monitoring the location of the objects themselves and of recording all updates to the live database on a regular basis.

One consolidated TMS for the Textile Collection

For more than a decade after the Ratti Center opened, the only way public visitors could access information about the textile collection was to schedule an ap-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

pointment and use TMS in the Center itself. This process was a bit cumbersome for researchers, many of whom were not aware of the extent of the collection and what might be of use for their projects. Once the export of data from TMS to the collection database on the museum's website had been effected, and the information had been consolidated into one main museum database, the public component of TMS was no longer supported by Gallery Systems, and thus no longer available to the public in the Center. By 2013, the Center started to use the online collection database via web access for outside visitors, and only museum staff still have access to the original search components.

The transition to one collections management system for the entire museum required the standardization and formatting of the main data entry fields, such as Constituent Names, Object Names, and Cultures, and of Location Sites and Sub-Sites. Component numbers and the different ways they were used within individual departments posed a problem for the Ratti Center, which tracks every accessioned object and every accessory related to an accessioned object. Most departments did not complete this tracking. During the transitional year of standardizing each department's components and accession numbers had to be updated to coincide with those in the Ratti Center's records.

One major concern was the level of access various staff members would have to confidential information and home location information. TMS has always had an extremely reliable security system that allows the assignment of a different range of access to each staff member. For instance, the Senior Manager of the Ratti Center may access all location and storage information for the objects stored in the Ratti Center, but that same Senior Manager does not have the same level of access to information for the Costume Institute's collection or for Drawings and Prints, for example, whose collection is not under the Ratti Center's jurisdiction. Accordingly, while the Senior Manager may change location records or storage information in TMS for the objects that are stored in the Center, she does not have the authority to see or change information for the collections belonging to departments not under her jurisdiction. The same concept applies to adding records or uploading images to individual records.

The diverse features offered by TMS, and its robust capabilities, make it an excellent program for The Met's needs. It can manage multiple collection details, and it is also the backbone for all loan agreement information, exhibition history, location history, and constituent information relating to an object's provenance, among other areas. TMS has withstood the test of time in The Met and proven itself to be an appropriate and efficient choice for the museum.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

Digital asset management systems at The Met

The museum maintains a separate Digital Asset Management (DAM) system, NetX, to archive and make accessible the more than one million master images of objects in its collections. The DAM also stores digitized historic images, visual records of special events, gallery views, construction photos, PDFs, audio files, and video. Staff can search images, download high resolution files, and share visual assets with colleagues. The DAM is closely integrated with TMS and delivers new images to the database nightly through an automated process. Object information from TMS is also sent to NetX every night through an additional process.

Maintaining a separate DAM allows staff to create multiple workflows that route and track assets. This permits them to manage the creation, retrieval, or updating of images throughout the imaging process. When a request for new object photography is made, the image is routed to the proper folder within NetX, deployed to TMS, and ultimately appears on the museum's various digital platforms, including the website, audio guide, gallery kiosks, and mobile app.

Case study: Database consolidation

The museum's cataloging department was started in 1910. During the early years cataloging staff typed object information onto catalog cards for all new acquisitions. By 1914 over eleven thousand catalog cards had been created for the collection. The museum's first cataloging manual was created at this time and remained in use for over fifty years. The cataloging department, which by the 1960s had grown to thirteen people, was disbanded in 1972, and each curatorial department took responsibility for cataloging its own collection. That change also marked the beginning of a movement away from museum-wide cataloging standards.

As a direct outgrowth of decentralized cataloging, when departments began using computer database systems to document their collections in the 1980s they selected their own systems based on their specific needs and requirements. Six different database programs were in use across the institution by 1995, which resulted in data silos accessible only within a single department. After the successful implementation of TMS in the Ratti Center, TMS was installed in each department gradually over the course of eight years. Existing databases were converted to TMS, while continuing to exist as separate systems. A major factor in the gradual installation of the system was museum infrastructure. The

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

Met is made up of a series of buildings and components, covering an area of two million square feet, with some sections over a hundred years old. Installation of network wiring, which had to be completed before a department could receive the new system, was a long and complex process. The museum eventually had twenty-three separate instances of TMS across seventeen departments, with some curatorial departments maintaining multiple databases for study or archival collections.

With the implementation of TMS museum-wide, The Met progressed to having a common collections management system used by all departments. A small collection information team of three people provided overall system support, user training, database maintenance, and report writing for staff. Maintaining separate database systems created several challenges, particularly because there were no institution-wide cataloging standards. Several departments created their own guidelines, but this still resulted in inconsistencies in the way fields were used, as well as in terminology, and name authorities since guidelines were specific to each department and not common across the museum. These inconsistencies came to light when object information was published on the museum's website as part of the online collection, since data for the website was taken directly from TMS. The data inconsistencies made searching difficult for both online users and internal staff. Administrative staff conducting museum-wide inventories had to search each separate database. Cross-collection searching by artist, donor, or time period was not possible.

Another result of the decentralized system was that database maintenance was burdensome. Work and preparation for system upgrades was multiplied due to the large number of databases. Testing routines had to be performed on each of the twenty-three databases. Major upgrade testing would occur over a period of months and require large amounts of dedicated staff time, as IT staff needed to schedule routine database backups and maintenance jobs for each database.

The decision to consolidate the separate databases into a single system was made in 2009, and the project took place over four years. It was a large and complex endeavor that involved collaboration between curatorial, IT, and collection information staff. The first phase was to establish institution-wide cataloging standards. With the Cataloging Cultural Objects (CCO) publication providing overall guidance, a committee made up of curatorial, library, and editorial staff was formed. The committee focused on four main clusters: Who (creators and cultures), What (object naming), Where (geography and nationality), and When (object dating). Priority was given to the fields that were most likely to be searched such as artist name, date, and object type. Guidelines included rules on data format, field use, and scope of content.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

Establishing standards that could be applied to the breadth of the museum's collection presented a challenge. Works in the collection cover a period of five thousand years, were produced on six continents, and are of an extensive variety including paintings, works on paper, musical instruments, decorative arts, costumes, household and ritual objects, and arms and armor. Some of the new rules required new data to be entered. Birth and death cities became compulsory for artists, along with nationality and life dates to provide more context for creators. In some cases, new rules for database field use required that departments change long-established data entry practices.

The work of the standards committee resulted in a written cataloging manual for the museum, the first since the original manual was introduced in 1910. The manual, which is constantly evolving, has acquired several supplements since it was first released. The supplements became necessary as the museum began to do more in-depth cataloging, acquired new types of works like time-based media, and expanded its holdings by Native American artists.

After the museum-cataloging manual was completed and documentation standards were put in place, the consolidation project moved into a data clean-up phase. It was important to clean-up as much data as possible prior to consolidation to ensure that the new single database would have the most accurate records possible for the most efficient and thorough searching and usability. Collection information staff conducted basic global cleanup tasks that addressed searchability, such as reformatting dates, moving data, and populating required fields. These adjustments affected hundreds of thousands of records, with over one million rows of data updated globally through automated methods. Other cleanup efforts focused on usability. Authority lists were evaluated, standardized, and refined. Administrative fields related to acquisitions and inventory were reviewed and documented in the cataloging manual to ensure consistency across departments. The storage location authority was revised to make it easier to identify to which departments each location belonged.

The greatest challenge in data cleanup was addressing the several thousand duplicate name authority entries across the twenty-three databases. Many of the same artist, donor, lender, and institution names were entered across departments, often with different spellings, life dates, and nationalities. These duplicates needed to be reconciled and made consistent before consolidation in order to produce accurate search results in the new database. Artist names were made a priority. Lists of names were sent to departments twice a month for review and comments. Any differences of opinion over this information among different departments were sent back to departments for resolution. Early lists required several rounds of curatorial review to finalize and complete.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

The process of identifying and standardizing artist names took two years to complete, with over four thousand names standardized. Several cataloging issues were resolved during the process, including formatting rules, the use of modern day place names, and the use of specific nationalities like British, Netherlandish, Flemish, and Dutch. Curatorial input during the cleanup process was critical. Artists could not be standardized without the expert knowledge of the museum's curators. The Getty's Union List of Artist Names (ULAN) and the Netherlands Institute for Art History (RKD) database were referenced in some instances, but for the majority of cases museum curators chose to determine their own preferences for artist names. Standardizing non-artist names belonging to donors, lenders, and other institutions is ongoing in 2016 and will take several years to complete.

Another area in which data had to be refined had to do with generalized information where specifics were necessary. Many records lacked data about the specific object. In such cases objects would be classified under a general object grouping such as furniture, sculpture, ceramics, or prints, and given an official artist title or one describing subject matter with no further details about the object type. This made searching for all tables, for example, or for all medals or figurines, or for specific types of ephemera problematic to both museum staff and users of the online collection. In order to improve indexing and searchability, title and object name became required fields. Though this requirement often resulted in duplicate information being entered in the title and object name fields, especially for decorative arts objects, the benefit of having a consistent field to use for searching was considered more important than dealing with the inevitable duplication.

A further complication was that standardizing terminology for object names was not part of the cleanup project. With over forty thousand unique object names in use across departments it was not feasible to standardize terms. Instead, the TMS application includes Art & Architecture Thesaurus (AAT) assisted search functionality which allows a variety of equivalent terms to be used when searching object records.

Prior to consolidation, departments would routinely enter objects belonging to other departments in their databases for the purposes of reference, inventory, and exhibition planning. This practice resulted in thousands of duplicate object records when the systems were consolidated. Once the new database was in place, staff in each department were tasked with reviewing data in both or all records to ensure that all relevant information was added to the canonical record. When this was completed the duplicate entry or entries were deleted. In many cases critical information such as location history and images had to be copied to the canonical record. In other cases data entered in one object record

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

differed from data entered in the other. This conflicting information included differences in attribution, titles, dimensions, and storage location, and required departments to review their records carefully and determine what information was accurate.

In addition to preparing the database content for consolidation, technical work was being done by IT staff. The process of merging twenty-three databases required developing several complex procedures to populate tables with values from each separate database, create new unique identifiers and primary keys, maintain database integrity, and ensure that integrated applications like the museum's digital asset management system, gallery kiosks, and the website continued to function properly and be linked to the new database. The IT team worked closely with the software vendor to understand the business requirements of the application. The entire consolidation process was systematized and re-systematized several times prior to launch. This was critical to the success of the project as each iteration exposed errors, bugs, and oversights that were able to be addressed and corrected over several months.

The launching of the consolidated database in July 2013 marked the first time the museum had a single database of its entire collection. The new system had over 350,000 records and nearly one million images. Data in the online collection appeared consistent and standard. Searches by artists, lenders, and donors across departments were possible. Lists of new museum acquisitions could quickly be generated. Curatorial staff now had access to records in other departments, which aided research and exhibition planning.

The new consolidated database had a major impact on the museum's registrar's office. With records from all departments centralized in a single system, registrar staff were able to use the database to track loans and shipments. Prior to this, registrars used separate bespoke systems to perform their work. The new database allowed them to take full advantage of the administrative functionality of TMS. New workflows were created to track and document outgoing loans, with details about borrowers, venues, exhibitions, insurance costs, and display requirements. All shipments in and out of the museum could be tracked in the consolidated system. Registrars use the system to record all shipping details, transit information, packing requirements, and courier instructions.

As a consequence of new types of administrative data being entered in the system, departments can generate a large number of reports from the database, saving invaluable staff time and creating more efficient workflows. Over one hundred official museum documents are now generated from TMS, including acquisition papers, board reports, exhibition checklists, loan agreements, insurance certificates, and packing lists. Each new workflow is fully documented, presented to staff, and added to the museum's cataloging manual.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

Implementing a single collection database was a major cultural shift for museum staff and transformed the way departments worked. After decades of cataloging objects in their own way, in their own systems, curatorial staff were now obliged to follow new rules and open their collection data to staff outside their departments. A key factor in the success of the project was the active support and mandate of the museum's director. Implementing any major institutional change would not have succeeded without executive urging and encouragement. The director publicly stated that a single database with museum-wide cataloging standards to better document the collection and provide greater access to our works for the public was an institutional necessity, and he made it a priority. This robust public statement underscored to museum staff the significance of the project, and helped increase cooperation between departments. Several years earlier in an attempt to establish museum cataloging standards, a team of consultants led several meetings with curatorial staff and produced a thorough written report outlining its recommendations. However, because there was no urgency to consolidate at the time, the project was not a museum priority and the recommendations were not implemented.

Constant and open communication with database users played a major role in the project. The goal of the project communication plan was to present no surprises or unanticipated concerns to curatorial staff at project launch. Throughout the four-year project regular monthly update meetings were held, additional meetings were called to discuss specific topics and resolutions, regular reminders were issued, and staff were given access to a test consolidated database to allow them to review their records and become familiar with the new system. Training sessions were also held for each department to help staff navigate the new database and review new cataloging rules. Being kept informed throughout all stages of the project helped prepare staff for the large-scale changes in their work. The project was conducted as a wholly collaborative effort between curatorial and collection information staff.

Active involvement of curatorial staff, both collections managers and curators, was also vital to the success of the project. As subject experts, curators were an important part of establishing cataloging standards. Equally important were the department collections managers who are the main database users and the ones who oversee data entry. Providing both groups a voice gave them a greater sense of ownership in the project and brought more validity to the end results.

For museum staff the consolidated database has become a vital tool for documenting research, planning exhibitions, conducting inventory, and tracking loans, and has saved inestimable staff time that is reflected in improved workflows. Use of the database has expanded to include more archival collections

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

and archaeological excavations. Departments have also begun to use the system in new ways to document information such as scholar visits and gallery events, activities not previously recorded in the database. For the public, the new system has had a significant impact on the way collection information is delivered to the museum's public platforms, providing a central repository for the most up-to-date information about the collection for the website, gallery kiosks, the mobile app, and audio guides. All these benefits continue to affirm the importance of cataloging, an ongoing process that will ensure records are up to date, accurate, and accessible.

Conclusion

Technology is pervasive in today's world and is relied upon for learning about various cultural institutions and the collections they house. For instance, visitor interaction with art works via new technology has begun to be studied closely due to the increased use of digitization to catalog entire collections. Information professionals such as librarians, archivists and collections managers are entrusted with assembling relevant data and organizing it in a retrievable manner, and they are trained to understand the dynamics of museum informatics and the ongoing advancements in technology. The field of museum informatics is somewhat peripheral to the field of library and information science, yet it is increasingly becoming intertwined with the responsibilities of the information professional who classifies and organizes information electronically in order to make it manageable for end users to find and utilize.

Among the information professional's most important roles in museums is that of curator of information resources. The future appeal of museums and all organizations of cultural heritage depends on the efforts of information professionals who recognize the challenges of museum informatics and the needs of libraries, archives and museums and also of visitors to such institutions. The implementation of information policies supports both LAMs and visitors, and allows for continuing collaborations among various parties, such as museum to museum, museum to school through outreach programs, and museum to scholar via an individual's research activities.

The experience of The Metropolitan Museum of Art was not unique but helped to establish a protocol for the cataloging of a collection on a collections management system. Beginning with the creation of an online catalog of the textile collection over twenty years ago, The Met has succeeded in uniting all

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

data on its extensive holdings in one robust system that can continue to develop as the collection grows.

References

Baca, Murtha, and Visual Resources Association. 2006. *Cataloging Cultural Objects: A Guide to Describing Cultural Works and Their Images*. Chicago: American Library Association. http://cco.vrafoundation.org/.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

Emily Griffin and Rachel Lipkin

10 How Metadata Informs Purpose: A Case Study of NYPL's Open Source GIS-Driven MapWarper Tool

Introduction

As cultural heritage institutions and governmental agencies have made their collections and services available online, digital maps have been capitalized upon to create and to convey complex information about global resources and trends. This is most likely because "images add a visually stimulating accent to any website, and images of maps evoke a certain curiosity" (Evans 2015, 7), and because maps enable "extremely flexible data handling" of statistics (Southall 2011, 149). Whether because of the inherent indexing value of maps, or because of the vast number of collections containing geographic information (called GI), a plethora of dynamic map tools are utilized by digital libraries, museums, and archives to encourage curiosity driven discovery (Favretto 2012, 162).

Geographic information systems (GIS) are "information system(s)...designed to store, retrieve, manipulate, update, analyze, display, and interpret spatially-related data" (Strasser 1995, 278). GIS was created in the 1960's and primarily used by experts analyzing the allocation of natural resources (Strasser 1995). GIS technology remained sequestered to use by government employees and users in specialized fields (Strasser 1995) until academic librarians and experts on K-12 education advocated for the creation of minimal GIS; software that facilitates meaningful GIS use without intimate knowledge of geology or statistics (Fitchett and Good 2012; Enoch and Gold 2013).

Key terms

"Rectify" refers to the process of aligning the map data over present day OpenStreetMap data. This is accomplished by identifying accurate control points that appear on the historical map and map provided by OpenStreetMap.

"Control points" are points added to a historical map and the corresponding OpenStreetMap of the area by users. If the control points don't match both

https://doi.org/10.1515/9783110337419-011

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

maps, the alignment of the map won't make sense, and the results will be inaccurate.

In the context of MapWarper, "Layers" are collections of rectified map images derived from maps that appear in the same atlas. Once a historic map is successfully rectified, the resulting image is saved as part of an atlas's layer. In many other GIS tools, each rectified map could be saved as its own layer. However, MapWarper does not support this function.

"Raster map" is a GIS industry term for a digital image of a map that cannot be edited in any way (e.g. a place name cannot be added or a "pin" cannot be dropped, as in Google Maps) (Ordnance Survey 2010). In MapWarper, the digitized images of maps are examples of raster maps.

"Slippy" is a term referring to web maps that allow users to change their view of a map by clicking and dragging their mouse.

The Case: NYPL's MapWarper portal

This chapter discusses a minimal GIS tool made available by the New York Public Library (NYPL) called the MapWarper portal. MapWarper is an open source GIS tool coded using Ruby on Rails. The NYPL created the portal to showcase their Firyal Map Collection. Over 10,000 of the available maps depict New York State and the surrounding area; maps range from the sixteenth to the nineteenth centuries (NYPL 2016b). The maps give patrons access to information about transportation lines, real estate development, industrial development, and topographical studies, among many other areas of interest (NYPL 2016b). The main purpose of MapWarper is to invite users to align points on historical maps with corresponding points on real-time maps, allowing data from the present day to be layered over historic maps. This chapter also examines issues surrounding how GIS tools are affected by metadata schemas, and how the data protocol utilized by the NYPL affects the usability of the portal.

Using MapWarper

The MapWarper portal is currently divided into five tabs that allow users to discover resources five different ways. The "browse all maps" tab facilitates the discovery of historic maps through a search bar capable of searching by text, image ID, or collections ID. Users can refine their searches by year, and limit

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

their search to rectified maps only. The "browse rectified maps" tab is similar to the "browse all maps" tab, but this tab only shows user-rectified maps. The "browse all layers" tab facilitates discovery of layers with the same layout and search bar options.

The "find maps by location" tab features a two-pane view: (1) the right-hand side shows a global "slippy" map that users can use to zoom in on a location, and (2) the left-hand pane shows maps associated with the place the user zooms in on. The "find layers by location" facilitates the discovery of layers with the same two-pane, "slippy" map layout of the "find maps by location" tab.

It should be noted that MapWarper's search function only pulls from the metadata related to a map's title. Unfortunately, not all resources have been given titles that consistently mention the location of the map. For example, a search of the term "Manhattan" will yield search results with titles similar to "Manhattan, V. 1, Plate No. 85 Map bounded by Willett St., E. Broadway, Suffolk St., Broome St.," however, maps of Manhattan with no mention of Manhattan in the atlas title or map title will not be listed as a result.

Literature review

Open source digital humanities projects

Open source is the practice of making the source code of software open to the public and freely available for use, modification, and redistribution. An interesting example of modern open source use is Tesla Motors. Elon Musk wrote and published a blog entry in 2014 explaining that, "We believe that applying the open source philosophy to our patents will strengthen rather than diminish Tesla's position" (Musk 2014, par. 1). After having better understood what open source entails it's much easier to then dive into what the digital humanities are. According to Opensource.com, "the digital humanities are where traditional humanities scholarship—or, the academic study of arts, language, history, and the like—meets the digital age. By using technology in new and innovative ways, digital humanities scholars can create research projects that explore topics in ways that were not possible (or were extremely laborious undertakings) before computers" (Holm 2013, par. 6). Open source and the digital humanities as an academic pursuit would not have been possible without the strong improvements on, and outcomes of, technology today.

With more and more public and private institutions absorbing digital resources and attaining new or improved hardware and software, digital human-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

ities stewardship now has a global reach. Text and data mining, digital publishing, and information retrieval are just a few key features that digital humanities research projects have accomplished. One example of a digital humanities project that utilizes open source code is the Kindred Britain Project funded by Stanford University Libraries.

Kindred Britain is a network of over 30,000 individuals, many of whom are iconic British celebrities and iconic figures, connected through family relationships of affiliation, marriage, or blood. Kindred Britain is a vision of a nation's history represented by a large family affair:

> What do they do: Kindred Britain takes the oldest form of network visualization – the family tree – and brings this diagram from the age of parchment into the midst of the new possibilities for knowledge and analysis enabled by 21st century digital technology. With around 30,000 individuals included, there are over 897 million different paths through the network. Many fascinating and defamiliarizing relationships emerge here that have remained hidden until now. And these connection-paths generate an almost infinite plenitude of local narratives. A few exemplary instances are spelled out in the 'Stories' section here: the rest are there to be extracted and told following your own interests. (Jenkins 2013, par. 1)

Kindred Britain and MapWarper engage users in a historical and cultural regard, which enables patrons and community members to feel connected to the platform. Kindred Britain and MapWarper have a wide variety of differences, yet they share some deep similarities in their mission and purpose. The key matters in these projects are that they strive to be (1) open source, (2) free educational resources, and (3) informative learning experiences. Whether they were created in a public library or in a private university setting is meaningless if these projects do not reach their stated mission to maintain the open source guidelines. MapWarper is a perfect example of an open source digital humanities project that is on the cusp of being successful while continually growing and developing.

Historians adopt GIS

GIS was designed in the 1960's as a result of experimentation with early computer software and the large amounts of available geographic data (Strasser 1995). The technology remained sequestered for use by government employees and any specific users with unique access to data sets and hardware appropriate for use with early GIS software (Strasser 1995). Academic librarians, as well as eager scholars who advocated for the use of GIS as a teaching tool, can be

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

credited with the widespread use of GIS in liberal arts and health sciences scholarship prevalent in contemporary academic literature (Fitchett and Good 2012; Enoch and Gold 2013).

After Google launched Google Maps in 2004, high-resolution images of the earth became easily accessible (Lamb and Johnson 2010; Fitchett and Good 2012). Subsequently, the use of GIS as education technology became commonplace in American institutions of higher learning and various educational settings worldwide (Beeson 2006; Shepherd 2009; Fitchett and Good 2012). Circa 2004, approximately twenty percent of all private learning institutions offered degrees in geography and coursework in GIS. A great number of universities, notably Harvard and Yale, developed intensive GIS learning centers to examine political phenomena such as genocide, diaspora, and colonial oppression (Weimer and Reehling 2006; Fitchett and Good 2012; Enoch and Gold 2013).

Open source GIS tools

The most popular GIS tool is the QGIS (formerly known as Quantum GIS) software package. As stated by gisgeography.com, (2016) "QGIS is jam-packed with hidden gems at your fingertips. Automate map production, process geospatial data, and generate drool-worthy cartographic figures. There's no other free mapping software on this list that lets you map like a rock star than QGIS" (par. 6). In addition to mapping, QGIS works with other companies, such as Halftone, to aid in restoring the grayscale texts on old or preserved maps. Another example is the Whitebox GAT project operated by the University of Guelph, Canada. This project began in 2009 and was developed as a replacement system for the Terrain Analysis System (TAS). The intention of the Whitebox GAT project was to have broader focus than its predecessor:

> Positioning it as an open-source desktop GIS and remote sensing software package for general applications of geospatial analysis and data visualization. Whitebox GAT is intended to provide a platform for advanced geospatial data analysis with applications in both environmental research and the geomatics industry more broadly. It was envisioned from the outset as providing an ideal platform for experimenting with novel geospatial analysis methods. Equally important is the project's goal of providing a tool that can be used for geomatics-based education. (Lindsay 2015, par. 2)

Open-access software, such as Whitebox GAT was designed to encourage end-users to examine the algorithms and details of its implementation associated with the software artifacts. Open-source and GIS systems strive to build interac-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

tive platforms with effective storytelling, creative user experiences, and code transparency. While these key components are present in NYPL's MapWarper project, a comparative analysis of the methods used by this project with other similar projects might prove interesting.

MapWarper versus Kindred Britain

MapWarper and Kindred Britain are both tools that are freely accessible for the general public, available online, and both require the patron to open an account. The difference between the MapWarper Project and the Kindred Britain Project is their origin. NYPL, an institution supported by public funding, created MapWarper whereas Kindred Britain was conceived by Stanford University, a private organization. MapWarper enables users to transpose historical maps onto present-day maps by rectifying them and to view historical maps. However, MapWarper does not organize their maps in a clear timeline, which can hinder users. On the other hand, Kindred Britain collects stories, genealogy, and geographic information and connects the information to a singular and fluid timeline. The organization and standardization of entries in Kindred Britain makes it easier for users to navigate the site and engage creatively with the resources provided.

Kindred Britain offers cultural and historical connectivity across generations. This connectivity is also found in MapWarper through historical maps. For example, MapWarper provides thorough and detailed maps from all over the world and from different time periods, allowing patrons of all ages and backgrounds to connect with and enrich historical resources. Although Kindred Britain and MapWarper are not considered social media outlets, they both create digital spaces for patrons to connect in a free, informative, and educational manner.

While Kindred Britain and MapWarper share some similarities in their pursuit of open source, there are some stark differences in MapWarper compared to other thriving GIS projects, such as, Geographic Resources Analysis Support System (GRASS) GIS and GeoServer. The authors of this chapter have chosen to use data from OpenHub to compare and contrast MapWarper with GRASS GIS and GeoServer.

Both GRASS GIS and GeoServer are popular GIS and open source systems used by the general public. MapWarper has a very low activity rate, which means MapWarper is accessed on fewer occasions when compared to GRASS GIS or GeoServer. High user activity is the goal in many open source systems

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

because it means the systems are engaging their audience, the community, and their patronage. Without high user activity MapWarper is not attaining their goal and therefore not growing because the lack of patronage and community participation is low. Another key difference is the estimated cost of the three projects. MapWarper's cost is relatively low compared to GRASS GIS and GeoServer, which can exceed into the millions of dollars. Lastly, the largest and most important line-item to focus on is the 'contributors' section. MapWarper has roughly eight developers contributing to the platform at any given time. Both GRASS GIS and GeoServer exceed seventy developers at any given time. This vast difference in contributors is what divides a successfully growing open source platform from a platform that stalls and remains to be antiquated.

Based on the information that was provided by OpenHub, there is an indication of a lack of participation in the MapWarper platform. There is more to say about the contrast between all three of these platforms, but based upon the limited information provided it can be concluded that MapWarper does not have the sufficient amount of contributors to make this project retain its patron use.

Metadata and MapWarper

Metadata about items held in the MapWarper portal includes (1) the title of the image, (2) information about the atlas from which the image was taken, (3) a permalink to all the layers related to the image, (4) what year the image depicts, (5) when the image was last modified, (6) the user who last modified the image, and (7) how many control points have been set (if the map has been rectified). Full catalog entries from NYPL's Digital Collection are linked to each map, which is formatted in the Metadata Object Description Schema (MODS) metadata schema (NYPL 2016a). The metadata object description schema (MODS) is (1) maintained by the U.S. Library of Congress, (2) expressed using eXtensible Markup Language (XML) code, (3) utilizing language-based tags (as opposed to numeric tags), and (4) based on the MARC21 record format (Larsgaard 2005; Hider et al. 2015). Some attractive features of MODS is that the schema is designed to carry selected data from existing MARC21 records yet also serve as a tool to create original records (Library of Congress 2016).

While the MapWarper portal itself does not contain full catalog information for items, the fact that MapWarper items are linked directly to MODS catalog entries lead Griffin and Lipkin to conclude that select MODS fields have been transferred to MapWarper. Additional information about any user activity asso-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

ciated with the items supplements the MODS fields. MapWarper's search bar allows users to search by text, image ID, or collection identification (ID). As mentioned earlier, text-search bars only search title metadata fields in the collection, limiting the results.

MODS supports the collection, preservation, access, and the creation of digital collections by converting MARC21 records specially formatted for the lending library environment into records ready to use for a variety of different purposes (Library of Congress 2016). The elements of MARC21 capture information about item creators, copyright, and unique identifiers; however, the schema does not capture any information about the quality of data or any information that a GIS user would need in order to download and share data. Another attractive feature of MODS is that extensions of the schema can be written; this enables users to create their own supplementary set of elements and/or fields to the schema (Hardy and Durante 2014).

Other metadata standards used by the GIS informatics community

Content Standard for Digital Spatial Metadata (CSDGM)

The CSDGM was created in 1994 by the U.S. Federal Geographic Data Committee (FGDC) and was a direct result of Executive Order 12906. The committee was organized because geospatial metadata was being utilized by a plethora of government agencies and because of the disparate nature of geospatial data and data protocol (Mathys and Boulos 2011). The goals of CSDGM were to discover what information typical users might need to decide whether data were relevant to their needs, and what information users might need to determine how to manipulate and share geospatial data (Larsgaard 2005). Additionally, the CSDGM was the first attempt at setting an international spatial data standard (Mathys and Boulos 2011). The most current version of the CSDGM was released in 1998.

The standard dictates that every record contain an abstract and summary and has dedicated sections (each with their own fields) describing (1) identification information, (2) data quality information, (3) spatial data organization information, (4) spatial reference information, (5) entity and attribute information, (6) distribution information, (7) metadata reference information, (8) cita-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

tion information, (9) time period information, and (10) contact information (Larsgaard 2005).

ISO standard 19115 "Geographic Information: Metadata" (ISO 19115).

The ISO standard 19115 'Geographic Information: Metadata' schema is a set of guidelines regarding the description of geospatial data based on others set by a committee of representatives from countries in continental Europe and the United Kingdom. The official regulations were ratified in 2003 (Mathys and Kamel Boulos 2011). The schema was designed to facilitate the searching and harvesting of geospatial metadata, espouses the free sharing of geographic data, allows different groups to create different application profiles and guidelines, and is part of twenty-eight related standards specific to geographic information (Vardakosta and Kapidakis 2013).

ISO 19115 has mandatory elements including (1) topic, (2) spatial extent, (3) point of contact for further information, (4) and date. Fields are usually optional. ISO 19115 standards also require: (1) metadata packages for metadata entity set information, (2) identification information, (3) constraint information, (4) data quality information, (5) maintenance information, (6) spatial representation information, (7) reference system information, (8) content (feature, attribution) information, (9) portrayal catalog information, (10) distribution information, (11) metadata extension information, and (12) application schema.

Dublin Core (DC)

Dublin Core became widely used circa 1995 and is a relatively simple schema, only listing fifteen elements, none of which are expressly required (Larsgaard 2005). DC allows users to adapt fields depending on the object or the institution. For example, the coverage field can be used to document geographic location or date. Additionally, users can choose between creating "qualified" records (records that contain values for all fifteen elements) or "unqualified" records (records that are missing values for one or more element). DC is widely used by digital libraries and interdisciplinary institutions seeking to digitize their collections; this is widely attributed to the fact that DC utilizes simple language and encourages flexibility (Vardakosta and Kapidakis 2013).

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

Key finding 1

Currently, there is a lack of consistent titles or naming of historical maps available through the MapWarper portal. Some maps have very long names that include the names of cartographers or titles of atlases from which the maps were found. Some map titles contain the names of New York City boroughs (the Bronx, Brooklyn, Manhattan, Queens and Staten Island), while others contain specific street names but no information about what borough the streets are in. It is recommended that the portal's search bar draw from several different metadata fields, not just the record or map title. Griffin and Lipkin also suggest that NYPL add a dedicated mandatory field for data about the borough the map depicts for every map. Alternatively, NYPL can work to systematically add borough information to the titles of all maps of New York City in the MapWarper portal.

As mentioned earlier, the Firyal collection houses maps from as early as the sixteenth century; because of a lack of precision of cartographic methods used in the past, certain maps available in MapWarper are simply not suitable to be aligned with present-day data (Favretto 2012). Furthermore, because of the inaccuracy of certain historical maps, users may be creating severely inaccurate rectified layers by attempting to match present day data with inaccurate raster maps (Favretto 2012). Consequently, Griffin and Lipkin recommend that NYPL add more robust data about the cartographic accuracy of the historical maps in their collection.

The benefits of adding metadata fields about the cartographic accuracy and whether the map is suitable for rectification are twofold: It is crucial that users be able to assess whether a given map will fit their needs (Larsgaard 2005). Additionally, tagging select maps as grossly inaccurate will aid in quality control by preventing the creation of inaccurate rectified layers (Favretto 2012). NYPL could convey the aforementioned crucial information by creating an extension to MODS based on FGDC's CSDGM sections describing Data Quality Information (to include information about the map' accuracy) and Spatial Data Organization Information (to indicate more relevant details about raster map scaling). It's also recommended that MapWarper adopt the importance that FGDC's CSDGM schema places on abstracts—this would facilitate easy browsing and assessment of maps while retaining a compact list view of results.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

Key finding 2

Presently, the NYPL facilitates the discovery of rectified layers through the "find layers by location" tab; users can open all 374 layers, or use a "slippy" map to find layers based on their location. Users may also use the "browse all layers" tab to find layers by text search or by collection ID. However, the results of a text search are hard to evaluate due to the lack of robust data about the layers found. There does not seem to be a consistent protocol regarding metadata creation at the layer level. For example, some layers are simply titled "Whole" or "Volume Two," while other layers are titled in the style of "Manuscript survey maps of various parts of New Jersey" or "Vermont." Some layers have data about the date they depict and the atlas from which they were taken, while other layers lack context due to a paucity of visible metadata.

The consistent application of a metadata schema to layers in MapWarper would significantly aid both users hoping to cite accurate data sets, and aid staff responsible for maintaining the portal. Favretto (2012) notes that researchers need to be aware of the accuracy of the resources they intend to incorporate into their studies, otherwise important studies can be skewed by the use of imprecise data (162). Mathys and Boulos (2011) aptly note that improperly documented layers are "likely to become redundant" (2) and that "datasets cannot easily be revealed to the wider geospatial community" (2). At present, it is unclear exactly how many layers of MapWarper are redundant and how the storage of redundant data impacts the loading time of the program.

Consistent language applied to titles of layers, with the addition of visible metadata about the dates covered, locations covered, and the historic context of the raster maps in MapWarper would resolve the barriers to the discovery of layers currently present in the portal. Moreover, the addition of the collection of metadata describing each rectified layer in MapWarper will also improve the usability of the portal.

Key finding 3

Charles Vickery wrote an article in 2015 for Opensource.com titled *How to make your open source project thrive*, which detailed the most basic techniques any open source platform can integrate for future success. The first step is to define the project's success. Success can mean a lot of different things for a lot of different people. If the MapWarper portal's success is defined as local patrons and

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

community organizers utilizing this tool, then the MapWarper can be considered successful.

However, if NYPL defines success as reaching a global audience, then NYPL must follow Vickery's (2015) third rule: recruit core contributors and market and promote the project. Vickery (2015) references a popular open source project, ssh-chat, which his colleague and close friend, Andrey, started. Vickery (2015) says, "Andrey started asking for help, asking for improvements, and finding ways to bring more people into the project. In order to help build interest, Andrey reached out to people on Twitter and offered free Geo programming lessons in exchange for opening pull requests. This overcame the initial inertia of getting a few people involved and interested in the project" (Vickery 2015, par. 4). Unfortunately, marketing and recruitment for the MapWarper project has not been prolific. Griffin and Lipkin recommend that NYPL begin writing a blog about the portal that will feature updates and a clear mission statement.

Lastly, Griffin and Lipkin suspect that limited funding prevents NYPL from employing a team of developers that is large enough to ensure premium levels of usability for a collection as large as NYPL's map collection. It's recommended that NYPL consider soliciting donations by adding a link to the MapWarper site that will encourage the popular "click, then donate" model.

Conclusion

The importance of open source resources goes beyond community building and promoting free access to software. The accessibility and openness of these resources allows best practices in software development to be established quickly, and encourages industry standards to be set by actual users. New York Public Library's MapWarper portal allows people, within the New York City communities and globally, to contribute their best work to the portal. Open source tools, such as the MapWarper, offer an opportunity for a large number of contributors, perhaps some of the best in the world, to build on top of prior quality code. An additional benefit of open source software is the free access to learn while contributing to a specific platform. Most software companies, educational resources, and online coding schools are far too expensive or time consuming for the average citizen to consider. NYPL's MapWarper portal reflects this mission to the core. It is a free, open source platform that allows patrons to contribute their findings to software and provides users access to historical maps of New York and an opportunity to become familiar with GIS.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

Metadata characterizes data; in short, metadata answers the who, what, when, where, why, and how of questions for users of that data. Accurate and robust metadata is crucial to retaining search function accuracy and positive user outcomes. Metadata is a reflection of the platform, the institution, and the collection and also preserves the collection for future users. Craig Mullins said it best in his blog:

> So the struggle today is trying to organize the cadre of repositories and attempting to rationalize all of the disparate locations and sources where metadata resides. Some shops are better than others at doing this. But I firmly believe that only those shops who understand the importance of their metadata can truly thrive as leaders... I mean, when you think about it, if you don't know your metadata, you don't know your data... and if you don't know your data, how can you conduct business effectively or efficiently? (Mullins 2010, par. 5)

Mullin's (2010) suggestion should be considered carefully by many organizations that are eager to achieve and attract many coders, users, and contributors. If the organization that is running or operating the open source platform does not understand or know their own data, then conducting effective business is virtually impossible. Collecting metadata, especially collecting accurate and consistent metadata, is critical to the overall functionality and effectiveness of the platform.

While the developers of MapWarper would be able to reach a larger audience by adopting strategies from the open source community; best practices in adapting library catalog metadata to support the requirements of a digital GIS portal can be used to help resolve usability issues surrounding barriers to discovery present in the portal. The MODS metadata schema can support the description of complicated geospatial, especially because the schema allows users to build extensions onto the basic schema (Hardy and Durante 2014; Hider et al. 2015).

Several researchers have described strategies used to automatically generate metadata (Mathys and Boulos 2011; Hardy and Durante 2014). NYPL could create a considerable amount of metadata without expending large funds to pay employees to add metadata to each record manually by utilizing XML coding to automatically generate: (1) data that differentiates historical raster maps from rectified layers, (2) data about whether or not maps are suitable for rectification, and (3) standard metadata about which borough each map depicts. Another efficient way to add more robust data would be to prompt users to tag the rectified maps they create with metadata about scaling, useful monuments or streets, and the accuracy of the control points they set. This model is used by

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

many digital humanities projects, and could result in users feeling more integral to the survival of MapWarper.

Lastly, many institutions choose to create their own custom metadata schemas to support their digital geospatial collections (Mathys and Boulos 2011; Hardy and Durante 2014; Hider et al. 2015). NYPL should consider investing in the time to create their own metadata schema to support their specific needs if creating extensions to their MODS metadata would be too time consuming or result in inefficient lines of code in the MapWarper portal.

References

Beeson, P. A. 2006. "Uncovering the Secrets behind the Successful Integration of GIS into the Core Curriculum." *International Research in Geographical & Environmental Education* 15 (3): 274–277. doi:10.2167/irgee196f.0.

Black Duck Software. 2014. "OpenHub: Compare Projects." https://www.openhub.net/.

Enoch, Jessica, and David A. Gold. 2013. "Seizing the Methodological Moment: The Digital Humanities and Historiography in Rhetoric and Composition." *College English* 76 (2): 105–114. http://search.proquest.com.jerome.stjohns.edu:81/docview/1445186343?accountid=14068.

Evans, Rachel. 2015. "Improving Web Discovery with Maps." *Computers in Libraries* 35 (7): 4–9.

Favretto, Andrea. 2012. "Georeferencing Historical Cartography: A Quality-Control Method." *Cartographica* 47 (3): 161–167. doi:10.3138/carto.47.3.1147.

Fitchett, P. G., and A. J. Good. 2012. "Teaching Genocide through GIS: A Transformative Approach." *Clearing House* 85 (3): 87–92. doi:10.1080/00098655.2011.628713.

GIS Geography. 2015. "13 Free GIS Software Options: Map the World in Open Source – GIS Geography." GIS Geography November 18. http://gisgeography.com/free-gis-software/.

Hardy, Darren and Kim Durante. 2014. "A Metadata Schema for Geospatial Resource Discovery Use Cases." *The Code{4}Lib Journal*, no. 25 (June), http://journal.code4lib.org/articles/9710.

Hider, Philip, Ying-Hsang Liu, Carole Gerts, Carla Daws, Barney Dalgarno, Sue Bennett, Barbara Spiller, Robert Parkes, Pat Knight, Pru Mitchell, Raylee Macaulay, and Lauren Carlson. 2015. "Developing a Schema for Describing the Contents of the Office for Learning and Teaching's Resource Library." *Australian Academic & Research Libraries* 46 (3): 151–163. doi:10.1080/00048623.2015.1030846.

Holm, Joshua Allen. 2015. "Open Source and Humanities in the Digital Age." Opensource.com. March 3. https://opensource.com/education/15/3/open-source-digital-humanities.

Jenkins, Nicholas, Elijah Meeks, and Scott Murray. 2013. "Kindred Britain." Kindred Britain. http://kindred.stanford.edu/#.

Lamb, A. and L. Johnson. 2010. "Virtual Expeditions: Google Earth, GIS, and Geovisualization Technologies in Teaching and Learning." *Teacher Librarian* 37 (3): 81–85. http://hdl.handle.net/1805/8651.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

Larsgaard, Mary Lynette. 2005. "Metaloging of Digital Geospatial Data." *The Cartographic Journal* 42 (3): 231–237. doi:10.1179/000870405X77183.
Library of Congress. 2016. "MODS: Uses and Features." http://www.loc.gov/standards/mods/mods-overview.html.
Lindsay, John. 2015. "Welcome to the Whitebox GAT Project." Whitebox GAT. http://www.uoguelph.ca/~hydrogeo/Whitebox/.
Mathys, Tony, and Maged N. Kamel Boulos. 2011. "Geospatial Resources for Supporting Data Standards, Guidance and Best Practice in Health Informatics." *BMC Research Notes* 4. doi:10.1186%2F1756–0500–4–19.
Mullins, Craig S. 2016. "Data and Technology Today." *Data and Technology Today*. May 30. https://datatechnologytoday.wordpress.com/.
Musk, Elon. 2014. "All Our Patent Are Belong To You." *Tesla Motors*. June 14. https://www.teslamotors.com/blog/all-our-patent-are-belong-you.
New York Public Library (NYPL). 2016a. "About Digital Collections." http://digitalcollections.nypl.org/about.
New York Public Library (NYPL). 2016b. "About the Map Division." http://www.nypl.org/about/divisions/map-division.
Shepherd, I. D. H. 2009. "From Geography Department to Business School: Strategies for Transplanting GIS Courses Between Disciplines." *Journal of Geography in Higher Education* 33: 28–45. doi:10.1080/03098260903034038.
Southall, Humphrey. 2011. "Rebuilding the Great Britain Historical GIS, Part 1: Building an Indefinitely Scalable Statistical Database." *Historical Methods* 44 (3): 149–159. doi:10.1080/01615440.2011.589774.
Strasser, Theresa C. 1995. "Desktop GIS in Libraries, Technology and Costs: A View from New York State." *Journal of Academic Librarianship* 21: 278–282. doi:10.1016/0099–1333(95)90008-X.
UK Ordnance Survey. 2010. "What's the Difference Between Raster and Vector Mapping?" The Ordnance Survey Blog, March 31. https://www.ordnancesurvey.co.uk/blog/2010/03/raster-vs-vector/.
Vardakosta, Iligenia, and Sarantos Kapidakis. 2013. "The New Trends for Librarians in Management of Geographic Information." *Procedia – Social and Behavioral Sciences* 73: 794–801. doi:10.1016/j.sbspro.2013.02.120.
Vickery, Charles. 2015. "Four Steps to Creating a Thriving Open Source Project." Opensource.com. May 26. https://opensource.com/life/15/5/4-steps-creating-thriving-open-source-project.
Weimer, Katherine. H., and Pete Reehling. 2006. "A New Model of Geographic Information Librarianship: Description, Curriculum and Program Proposal." *Journal of Education for Library & Information Science* 47 (4): 291–302. doi:10.2307/40323822.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

Travis L. Wagner

11 Transcribe as Seen: Challenging RDA Regarding Gender in Moving Image Materials

Introduction

Resource Description and Access (RDA) offers more free-formed approaches to constructing metadata, decentralizing the need for rigid comprehension of specificity when providing expansive metadata to objects. It allows for more operability and nuanced ways of describing items with contextual differences in larger collections, while maintaining connectivity between collections. One can imagine the nuance placed on topics of identity within a given framework, particularly concerning self-defined identities. Yet, early iterations RDA failed to facilitate language to discuss gender; assuming a categorical binary, forcing catalogers to label gender as either male or female, thus completely disregarding how a person identifies on a spectrum outside this binary. This limitation proves detrimental to users seeking out queer and non-gender conforming materials. Through critical advocacy, this guideline changed to state that catalogers should provide gender information with an understanding that "gender is the gender with which a person identifies" (RDA Toolkit 2017, under "RDA Principle 9.7"). While one may read this as progress, it is progress built on avoiding problems of normative gender presumptions. Stating that the person must identify explicitly as a gender limits potentials of identity and burdens those claiming identities outside the limits of a male/female binary. Furthermore, the amount of materials which do not provide said statement are limitless, as evidenced by the amount of queer materials housed in moving image archives.

This chapter makes the case that moving image materials push the limitations of description within RDA's newest instruction for discussing gender. This chapter will use examples from archival moving image materials to show how a 'burden of proof' model fails to make visible both open statements of gender, as well as implied gender identities. Simply put, some moving image archival materials do not express the gender of their many subjects, and historical shifts in understandings of queer identity and potentially transgender representations in the moving image have altered how one discusses such encounters. This chapter critiques this function of RDA, while imagining alternative ways in which

https://doi.org/10.1515/9783110337419-012

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

RDA 9.7 could be expanded to discuss potentialities of gender, as opposed to avoiding the discussion outright.

Resource Description and Access (RDA) emerged within information organization practices as a promising and proactive structure to harness the near infinite variety of materials housed within cultural institutions. Advocates for the new description tool see the implementation as providing crucial depth needed for catalog records, databases and other information repositories. More important than the detailed nuancing available within RDA is the suggestion that, as an organizational method, RDA will be far more user-friendly, malleable and open to iterative improvements than preceding standards. This is most notable concerning RDA's predecessor and information organization standard the Anglo-American Cataloging Rules (AARC2). Furthermore, the implementation of RDA asserts not a dismissal of the place of MAchine-Readable Cataloging (MARC) records within information organization, but merely serves to invite expansion and afford evolution of descriptions in relation to previous points of limitation. Indeed, the focus of RDA is not to provide catalogers with a rigid set of rules to follow to exactitude, but alternatively a series of guidelines and suggestions for how to structure records. Such guidelines are deeply beneficial as they direct information organization professionals in a path towards cohesion and logic for materials that they may have never previously encountered within an organizational context. An example of these expanded guidelines include the adding and expanding of content types for resources. Specifically, these content types differentiate format to clarify that two items of information may have similar content, but are housed on divergent formats. One such example being the difference between a "cartographic dataset" housed on microfilm and one housed on a computer file. RDA suggests further nuancing in terms of how to describe these content types. Continuing with cartographic works, the newest guidelines for RDA offers helpful guidance in how to order longitudinal and latitudinal information for a map, while also suggesting the provision of both vertical and horizontal scale if possible (El-Sherbini 2013). While cartography specific libraries and archives might further nuance their records to expand descriptive elements of maps, this small example shows the directorial nature of RDA as it pertains to the generalist cataloger. It is but one of many and speaks to the quantifiable and easily identifiable components of RDA-infused information organization. However, not all components of the guided description are as productive and beneficial as the aforementioned example.

In an earlier iteration of RDA, users possessed an option for a line in which they could provide a gender for the creator of and persons represented within any information item. Located at 9.7 within the RDA Principles, the guide initially stated that a gender description can exist between "male," "female" and

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

"not known" and that such choices should come from whatever information might be available. In more current iterations, RDA has allowed for more expansive descriptions for gender that allow a person to "choose the gender with which a person identifies" (RDA Toolkit 2017, under "RDA Principle 9.7"). In this need to choose a gender, many criticisms have emerged, specifically related to the problems of presumed fixed gender identities, as well as the potentiality for multiple, fluctuating gender identities. Using a queer and feminist framework of analysis, cataloging scholars Amber Billey, Emily Drabinski and K.R. Roberto (2014) suggest that to force catalogers to engage in these practices delegitimizes many groups and specifically, silences those whose gender representation is fluid, purposefully complex and could be contradicted within the broad scope of "any sources" from which RDA suggests a cataloger can pull information (RDA Toolkit 2017). Specifically, Billey, Drabinski, and Roberto (2014) posit "gender and sex are always negotiated and socially constructed" and "fixing them, as RDA asks" results in such negotiations going unacknowledged (414). In no small way, RDA implores the cataloger to engage in a major contradiction, by first asking that said cataloger adhere to the gender with which the potential person identifies, but then to use any resource to affirm that gender. This includes resources, which may disavow said person's gender, as quickly as it is inclined to give it validation.

This chapter builds on the aforementioned work of Billey, Drabinski, and Roberto (2014) to challenge the perceived functionality of gender within RDA cataloging practices, exploring how such troubling identification practices become exacerbated within moving image archival information organization. By arguing that the moving image, within an archival context, is prone to a particularly high degree of interpretive work during cataloging practices, the chapter will nuance the trouble of misgendering, naming and defining identity as they relate to RDA guidelines. For such an interpretation to work, the chapter will borrow examples of moving images from the University of South Carolina's Moving Image Research Collections (MIRC) which could potentially trouble understandings of single, fixed gender identities, while also showing the multiplicity of ways in which gender could be iterative even within a given set of frames, never mind changes over years and decades. While reviewing the fixed gendered ties of the new RDA rule, discussion about the challenges of naming a person will arise, specifically as this too reflects potential challenges to gender identity. Provided in this section will be mock RDA records that explore these challenges, while also questioning if alternatives can, or should, exist. Finally, the chapter will make the case that while defining identity in an inclusive and non-restrictive manner within information organization is crucial, the real work needs to occur outside of catalog records and within interactions between users,

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

materials, and their respective descriptions, as to assure gender inclusivity exists beyond the confines of handful of MARC records.

As Billey, Drabinski, and Roberto (2014) point out within the conclusion of their critique of RDA, "binary is a central organizing feature of contemporary life, but it need not always be so" (420). Nonetheless, organizational structures emerge forcefully within information organization, particularly as a means to create distributable and retrievable information. Yet, those who define how these organizational systems can and should look often cannot detach themselves from a specific worldview when describing such content and emerging binaries and divisions may enact more harm than benefit. Indeed, in the pioneering critique of one of the most widely used of information organization standards, Library of Congress Subject Headings (LCSH), Berman (1993) suggests that they are a product of "parochial" and "jingoistic" longings by predominantly Anglo-American cisgender, white men (15). In this approach, Berman (1993) called attention to the ways in which information was organized to fit idealized users, who invariably reflected, or were presumed to reflect, those creating points of access. Building on this critique, Olson (2008) asserts that within such binary structures a certain degree of otherness is observed as a fixture of not being identified within a socially privileged self. Specifically, persons that do not reflect the aforementioned parts of Berman's (1993) critique become named, but the sameness presumed is not called attention to in any manner (Olson 2008). An example of this occurs within the Library of Congress where the presence of a heading for WOMEN COLLEGE TEACHERS, BLACK, without an equal heading for MEN COLLEGE TEACHERS, WHITE. A high degree of presumption occurs in how descriptive choices emerge and the lack of change to such discrepancies is telling to the field of information organization. As such, it is hard to imagine a careful and inclusive approach to gender description emerging, when the creation and implementation of binaries occurred by persons identifying within a cisgender binary. Not to say that this is impossible, but it is decidedly irregular. Furthermore, understanding that gender is not necessarily fixed between two points of masculinity and femininity is crucial to queer identity and the challenges this puts forth for such organizational approaches are worthwhile considerations.

Gender is a social construct that is situationally and historically iterative. One also changes and performs their gender, as a given moment requires. At times, this performance of gender can work to subvert notions of normalized binary gender identities, by moving to excessive and expressive versions of gender, as is the case with drag performances, while at other times existing between notions of femininity and masculinity that blur the purportedly fixed points between each. Gender is constantly at odds with social scripts for how

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

one must look within a certain 'normative' gender, resulting in scripts, which unjustly tie one's sexual identity to one's gender. As gender theorist Butler (1988) notes, "performing one's gender wrong initiates a set of punishments both obvious and indirect, and performing it well provides the reassurance that there is an essentialism of gender identity after all" (528). Simply put, one is made to feel unwelcome should they choose to perform a gender identity that does not 'match' their sex assigned at birth, whereas those who perform a gender that matches their sex assigned at birth (i.e. cisgender persons) are afforded the privileges of comfortably existing in society within both their sex and their gender. Gender then is not as a linear movement between two fixed points on a binary line, with male and female on opposing ends, but instead as a spectrum. This spectrum affords persons "movement between fluidity and categori[z]ation" (Monro 2005, 19). A resulting problem arises when cisgender persons attempt to place queer or transgender persons within constructs that do not acknowledge such a spectrum, relying instead on the assumption that any given gender identity is at the very least moving between masculinity and femininity. For queer identities and non-gender normative performances, this may not be the case as the person in question may reject both binary constructs, or even potentially occupy both simultaneously. To apply a defined label to any of these performances would be to silence the existence of the individual gender identity, and even placing iterative time gaps for each moment of gender identity as the MARC 375 field asks one to do, does not accept the fluidity of such gender. Further, when it does, even then it assumes segments of gender performance that can be comfortably and cohesively organized. As Drabinski (2013) argues of queerness and cataloging as a whole, "queer theoretical perspectives on classification and cataloging challenge the idea that a stable, universal, objective knowledge organization system could even exist; there is no such thing if categories and names are always contingent and in motion" (104). For Drabinski (2013) and most queer theorists within and outside of information organization, the need to name gender is ultimately anti-theoretical and misses the point of gender as a function of social performance. For many, gender is not a fluid thing and they exist statically within their respective identities and do not perceive this fluidity, however, for others this temporality is crucial to existence and relegation to such a fixed state, even posthumously, denies that person their humanity. Before moving towards how named gender identities are particularly a challenge to moving image archival materials with a potential queerness, it is crucial to point out how such naming actualizes a delegitimization of identity within a queer existence.

Jack Halberstam is a prolific scholar within queer theory, with his texts *In a Queer Time and Place* (2005), *The Queer Art of Failure* (2011), and *Gaga Femi-*

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

nism (2012) serving as germinal texts in a now prominent field of queer studies. Notably her works *In a Queer Time and Place* and *The Queer Art of Failure* circulate with the author as Judith Halberstam. This is because Jack Halberstam chooses to use the name Jack to represent their identity, potentially in relation to a queer transgender identity that does not accept the duality of gender and fluctuates between multiple identities. In terms of their gender, Halberstam (2012) has gone on record as stating that she prefers to use "floating gender pronouns" and further "refus[es] to resolve his 'gender ambiguity'" (under section titled "On Pronouns"). They have refused to "resolve" concerns over proper pronoun use and continue to do so in their writing and work (Halberstam 2012). So if one were to use RDA's guidance of whatever gender the individual so defines, as allocated by whatever source one obtains, it would be plausible for a cataloger to choose their gender as 'male', 'female' and potentially 'not known'. However, Halberstam (2012) makes it clear that he would prefer it to be all and none of those simultaneously. This troubling of gender categorization speaks volumes to what Billey, Drabinski, and Roberto (2014) warn against within their critique of RDA's 9.7 standard.

Furthermore, to adhere to a MARC-infused version of RDA 9.7, the respective genders should include dates that correspond to when Halberstam (2012) was 'of that' gender, completely disregarding the mobility and multiplicity of potential genders. It is not the role of catalogers to attempt to define such identities as this can be difficult to do with any degree of certainty or cohesion, expressly when the individual discussed identifies outside the scope of a singular, fixed gender. From a queer humanist standpoint, it is an injustice to transgender, queer, and non-gender conforming persons to approach description from this manner. For those less concerned with advocacy, in a neoliberalist framework, it is hardly a cost-effective approach to information organization to spend time covering the evolving scope of the gender spectrum and it would be economically more worthwhile to simply do away with the practice full scale. Simply put, no benefit comes from the need to name gender within persons attached to items of information, for doing so presumes certainty and rigidity of gender identity. This is decidedly true for content which has a known, living creator, but the problem of naming gender as a descriptive item is exacerbated when dealing with items whose participants possess no informational aids or resources to help 'assure' their respective genders.

The Moving Image Research Collections (MIRC) at The University of South Carolina represent one of the biggest audiovisual archives within the Southeastern United States. While its collections range in type, the archive is most prominently known for their Fox Movietone Newsreel Collection and Chinese Film Collection. Like other moving image archives, MIRC also holds home movies,

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

regional films and educational films from film's inception and onward. Similar to traditional archival print holdings, MIRC obtains acquisitions in a myriad of ways, often arriving with little or no information. Unlike print materials, however, moving images are harder to immediately identify and often require specialized knowledge of fashion, engineering, automotives or film stock/video formats to pinpoint general years for a given moving image. Furthermore, moving image materials take a considerable degree of interpretive work beyond format descriptions and content length. Indeed with moving image archiving, many institutions simply describe things at an item or collection level, understanding that "researchers can better study individual items when each is examined as emerging from the larger context of the whole" (Leigh 2006, 37). Something like a series of home movies are given a donor name when possible and obvious portions of scenes receive a title line, but beyond this any interpretive work remains up to those working with the materials beyond their description and access point.

Deeper exploration of the records includes general notes about each and named entities when possible, but any iteration of unidentified persons receives the general name of 'people'. A similar descriptive process happens with both educational films and newsreel footage. Both often lack named entities since credits rarely exist for such works. Accordingly, the descriptive elements for these types of works generally describe the plot or events shown and only provide named access points when directly stated within the respective moving images runtime. Clearly, such methods of cataloging are not ideal for full-scale access, but without knowledge of the entities beyond their mere presence in a frame of footage, describing anything beyond generalities would be to enact unwarranted interpretive work, some of which could prove to be incorrect or exclusionary based on the potential misinterpretations or presumptions of the cataloger in question. More concisely, moving-image materials possess too many uncertainties to describe beyond generalities and, in such cases, less is arguably more. Keeping this focus on limited description as it pertains to moving image materials; it is here where the 9.7 standard within RDA is expressly a suggested practice in need of contestation.

An example of a moving image item held at MIRC is the *Phi Eta Club Musical Show – Outtakes* (1927). Part of the Fox Movietone Newsreel Collection, the footage shows fraternity brothers engaging in a drag show that involves various games of leapfrog and laughing. From 1927, it is easy enough to pass off the performance as parody and think nothing of it as a work beyond this. Yet, the film includes a section in which one of the men in drag kisses another man not wearing drag. An individual emerges from off-of-screen and attempts to break up the encounter (presumably believing the man to be attacking a woman), but

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

upon realizing that the other woman is a man in drag he steps away and allows the two to return to their encounter.

Setting aside the complicated nature of drag as a historical act, one can begin to wonder about the potential implications for a piece of footage such as this. Could the kiss be a part of the larger drag show or could it be indicative of a non-heterosexual desire, which would have been difficult to express openly in the 1920s? Arguably, one could make the case for either, but doing so would go against the role of the cataloger, who is limited to generalized statements about the film, which should leave interpretation open to the future users of the material. This caveat is acceptable for a moving image such as this; however, this concerns only the potential sexual orientations at play in Phi Eta Club Musical Show – Outtakes (1927). This says nothing of the potential of gender within the film. After all, sexual orientation and gender are two distinctly different components of a person's identity. The former represents the person's to which one is attracted (physically, romantically and otherwise) wherein gender is, as noted earlier, how one performs their identity within a series of socially constructed understandings of what it means to be masculine, feminine or non-gender conforming. Similarly, to name one's gender in a moment like this is not to also name their sex assigned at birth. To assume that all of the individuals in Phi Eta Club Musical Show – Outtakes (1927) are male's is as equally detrimental to saying that they are performing as the gender of men within the short clip. These things remain unknown, and further, in the case of this film, the perceived possibilities of gender shift markedly. For instance, the film concludes with an individual sitting at a table while applying the makeup and hair for a presumed drag performance. Caught in the moment of transition the figure could exist liminally between male or female, but this tension should not be presumed to be an eventual transition between one or the other, indeed, it could be either or both simultaneously. As shown above, even writing about a moving image like Phi Eta Club Musical Show – Outtakes (1927), without leaning on gendered pronouns is difficult and arguably a product of the way in which drag shows are discussed. Yet, if one were to follow the guidance that moving images necessitate generalizability, describing this entire scene with any gendered pronouns whatsoever would be detrimental to usability and researchability. RDA 9.7 as a functional standard that theoretically added onto the description of this piece of content would only further legitimize a need for pronouns in moments where they are not only unnecessary, but also inherently exclusionary. Take for example a version of RDA gender descriptions that borrows from the only available information (as suggested per guidelines) for Phi Eta Club Musical Show – Outtakes (1927), the original description for the footage lifted from camera operator's notes circa 1927. The description states, "Harvard boys borrow their sister's

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

clothes for female roles in Pi Eta Club musical show" (Moving Image Research Digital Video Repository 2013). Accordingly, one would have to adhere to the suggested gender component and only provide a 375 line with "Male," negating any offering that could suggest a complexity of gender.

Yet, simply correcting this hindrance by suggesting that there may be a transgender representation does not suffice. It bottles the potential transgender identity into a specific umbrella that fails to nuance the many ways in which one could be transgender, queer or otherwise. Again, this is not to say definitively that any of the persons in Phi Eta Club Musical Show – Outtakes (1927) are indeed transgender, but only that the presumption of maleness does a disservice to its potential alternative interpretations. However, here one would need to have repetitive iterations to reflect the many temporal and spatial shifts of the film, meaning that the already drawn out record extends multiple times in the case of Phi Eta Club Musical Show – Outtakes (1927). While all of those interpretations could be present within the work, accounting for them could result in potential misnaming or presumptions that such gender identities apply to specific persons, while not holding true for others. This stands in marked contrast to something like crowdsourcing or folksonomies wherein populations receive access to group-defined describing and interpreting. RDA, as it stands, works to implement additions to language practiced by a handful of institutional catalogers at most. While these catalogers may well represent many spectrums within representations of gender they do not represent every gender and are, as such, incapable of properly enacting or naming such interpretations. Furthermore, even if such genders are diversely included in the naming process, it cannot affirm with any certainty the identity of those within a moving image work. Phi Eta Club Musical Show – Outtakes (1927) is hardly an exclusive example of this challenge. Within the Fox Movietone collection alone exists works which purports to depict an 'all women' wedding with some women dressed in male drag. Again, it crosses boundaries of sexual orientation while simultaneously complicating notions of gender and sex. Outside of the confines of MIRC as an institution, similar issues arise within gendered representations on moving images. Albert Steg and Robert Vaszari (2014) discussed their challenges when discovering a home movie they colloquially titled "The Transformation of Crystal." In the film, what appears to be a male enters a room and slowly begins applying makeup and a wig and the image fades and cuts back to what appears to be a woman named Crystal. Though both were able to find information about Crystal as a popular drag performer, nothing is known about how the individual shown defines their own gender and, accordingly, no such measures should be taken within descriptive moves, RDA included (Steg and Vaszari 2014).

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

Crucially, it is also necessary to acknowledge that gender matters for how one's name is evoked within a catalog record, so merely changing the practice of gender identity does not exclude other actions from being necessary. An example of this emerges with regards to the many appearances of Joe Carstairs within the Fox Movietone Newsreel collections (Summerscale 2012). A prominent female boat-racer, Carstairs is most known for her speed in the water. Yet catalog records have referred to her as everything from Betty Carstairs to her name given at birth of Barbara Carstairs (Summerscale 2012). Likely due to cataloging research, which would have reified her name as her birth name, it was not until later in life that Carstairs directly demanded that she be called Joe (Summerscale 2012). Though never expressing her gender as anything but a woman, she did state that she "was never a little girl" and that she "came out of the womb queer" (Summerscale 2012, 14). Again calling Carstairs by any gender within the record would be overstepping interpretation given her clearly complicated relationship with gender, but how she decided to identify by name is hardly beyond descriptive alteration and shows a moment of active confrontation within cataloging that could serve to confront failures with regards to identity and one's gender.

While critiquing acts of gendering within RDA is a relatively easy undertaking, doing so does not negate the importance of gender within information organization. A person's gender matters and how this is talked about is not a thing to be dismissed as 'too difficult' to be approached. It is hard to know exactly why the creators of RDA felt it necessary to include such a standard within their organization methodology. In all likelihood, it was an aim at showing progress and inclusivity within institutional work, a way of saying that indeed gender does matter. Yet, it is expressly in this non-tangible manner where well-meaning attempts at inclusivity fail. To borrow from Sara Ahmed's (2012) work on diversity, such "work becomes about generating the 'right image' while failing ultimately in 'correcting the wrong one'" (34). Adding gender to descriptive work cannot happen purely because it is what you are 'supposed to do'. Focus instead should work beyond the catalog record to assure that a cohesive discourse around gender and inclusivity exists. Discussions should arise at the reference desk, in staff meetings and at conferences about how gender works beyond labels and how institutions can work to be more inclusive. Knowing that problems exist within texts and naming should not mean that the collective of information organization works towards finding complete solutions to any given challenge, particularly if the group that causes the challenge wants to deliberately exist beyond fixed boundaries and rigid definitions.

Knowledge about the failings of language and the failings of organizing identities can facilitate an "emergence of consciousness" and evoke a "critical

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

intervention in reality" (Drabinski 2013, 204). Moving image archival materials are a blatant reminder that descriptive analysis is simply not a subjective act, one has to interpret what is occurring and that interpretation is always situational. Similarly, interpreting and enacting gender is contextual and no single truth on what gender is exists beyond the decision of the individual embodying said gender. Knowing this means that information organization must step beyond the role of practitioner and be a social agent. Discussions about gender should be constantly iterative and at no moment presume to have provided a fixed answer. The presence of users may radically change the understanding of a piece of moving image material and its gendered potentialities. This is not a detriment to the role of the information organization profession. Instead, it should be a welcomed aid in the process of knowledge building. One might imagine a future version of RDA 9.7 functioning as a user-tagging space, where one adds and links, but never provides a definitive gender for the individual in question. Approaching the challenges of embodied and lived identities within the organization of knowledge requires curiosity, but it equally requires accepting moments of failure and the discomfort that comes along with such actions. When it comes to defining gender within RDA 9.7, the act has absolutely failed and accepting this may cause some to feel uncomfortable. This discomfort should serve as a moment of reflection and a chance to progress forward. The labor of curious persons continues to fuel the field of cataloging and information organization. It is time to embrace this curiosity, while always remembering to expect discomfort along the way.

References

Ahmed, Sara. 2012. *On Being Included: Racism and Diversity in Institutional Life*. Durham: Duke University Press.

Berman, Sanford. 1993. *Prejudices and Antipathies: A Tract on the LC Subject Heads Concerning People*. Jefferson, N.C: McFarland.

Billey, Amber, Emily Drabinski, and K. R. Roberto. 2014. "What's Gender Got to Do with It? A Critique of RDA 9.7." *Cataloging & Classification Quarterly* 52 (4): 412–421. doi:10.1080/01639374.2014.882465.

Butler, Judith. 1988. "Performative Acts and Gender Constitution: An Essay in Phenomenology and Feminist Theory." *Theater Journal* 40 (4): 519–531. http://www.jstor.org/stable/3207893?origin=JSTOR-pdf.

Drabinski, Emily. 2013. "Queering the Catalog: Queer Theory and the Politics of Correction." *The Library* 83 (2).

El-Sherbini, Magda. 2013. *RDA: Strategies for Implementation*. Chicago: American Library Association.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

Halberstam, Jack. 2012. "On Pronouns." *Jack Halberstam: Gaga Feminism and Queer Failure*. http://www.jackhalberstam.com/on-pronouns.

Leigh, Andrea. 2006. "Context! Context! Context! Describing Moving Images at the Collection Level." *The Moving Image* 6 (1): 33–65. doi:10.1353/mov.2006.0014.

Monro, Surya. 2005. "Beyond Male and Female: Poststructuralism and the Spectrum of Gender." *International Journal of Transgenderism* 8 (1): 3–22. doi:10.1300/J485v08n01_02.

Moving Image Research Collections: Digital Video Repository. 1927. "Phi Eta Club Musical Show – Outtakes." http://mirc.sc.edu/islandora/object/usc:14223.

Olson, Hope A. 2008. "How We Construct Subjects: A Feminist Analysis." *Library Trends* 56 (2): 509–541.

RDA Toolkit: Resource, Description, and Access. "Getting Started with the RDA Toolkit." https://access.rdatoolkit.org/.

Steg, Albert and Robert Vaszari. 2014. "The Transformation of Crystal." At *Bastards Film Encounter*. Raleigh, NC. http://bastardfilmencounter.com/more-info/.

Summerscale, Kate. 2012. *The Queen of Whale Cay: The Extraordinary Story of 'Joe' Carstairs, the Fastest Woman on Water*. London: A&C Black.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

Sonia Yaco, Saleha Rizvi, and Arkalgud Ramaprasad

12 Cultural Heritage Curriculum Crosswalk: Using Metadata to Connect to Curriculum

Introduction

Can our metadata work harder if we create a new discovery paradigm that uses them to connect collections to curriculum? At a time of shrinking resources in higher education for cultural heritage repositories such as libraries, archives, and museums, we also face increasing pressure to show that our services and holdings improve student-learning outcomes. This chapter describes a successful pilot program that created an automated system to recommend manuscript collections to be used in courses offered at the University of Illinois at Chicago (UIC). The methodology involved (1) collecting course metadata, (2) assigning authority-controlled subjects to courses, (3) creating recommendations of relevant manuscript collections for the courses, and (4) using the pilot recommendations in bibliographic instruction. Among the challenges of the pilot were the existence of a multitude of collection and curricular data silos. The study led to changes in library cataloging practices and prompted several initiatives that will make the curricular data more accessible to the campus community.

The success of the pilot suggests that the crosswalk could be useful to other institutions. However, the software created for the pilot required intensive manual data manipulation and is not scalable. To develop a more systematic approach, UIC researchers created an ontology, a structural definition, of the crosswalk. Having an ontology helps cultural heritage professionals, university administrators and systems decision-makers to speak a common language and reduce ambiguity to set desired outcomes. The cultural heritage professionals may speak of the different types of collections; the administrators of funding, publicity, and usage; and the systems developers of databases and knowledge management. While each may use terms particular to their domain, the integration of the three will be central to an effective solution. The ontology will help in the development of a software solution to automate matching, which will allow collection metadata exploitation to enhance the value of collections for a curriculum and will allow curriculum to drive the utilization of the collections.

This chapter begins with a review of literature and includes a case study, an ontology developed from lessons learned, and steps for turning the ontology

https://doi.org/10.1515/9783110337419-013

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

into practical solutions. While the pilot focused on manuscript collections, the lessons learned are applicable to all types of cultural heritage collections.

Linking curriculum to collections

Cultural heritage collections such as archives have played a key role in the intellectual life of universities since ancient times, yet schema to organize them and policies to allow open access to collections are relatively modern concepts. Melvil Dewey published the first uniform classification system for books in the United States in 1876 at a time when academic libraries began permitting faculty and students to use their collections actively (Weiner 2005). Some 160 years later, many universities as well as libraries, archives, and museums (LAMs), abound with metadata for both the cultural heritage and curricular realms in the form of course descriptions and repository catalogs. Yet there has been little effort to link these two sources. The current literature almost exclusively promotes face-to-face methods to promote the use of library resources in the classroom and in research.

The most common methods for integrating collections into curriculum are creating orientation sessions for undergraduates (Allison 2005), suggesting relevant collections and possible class assignments to individual faculty (Robyns 2001; Maher 2001; Rockenbach 2011), and collaborating with instructional faculty to develop courses (Kennedy 2006). A recent trend is embedding librarians into academic programs (Kesselman and Watstein 2009; Kvenild and Calkins 2011; Montoya 2013) to provide individual attention to students and faculty.

Some librarians have sought to build systematic links between collections and courses. In the late 1960s, William E. McGrath and Norma Durand (1969) assigned Library of Congress call numbers to all courses in the University of Southeastern Louisiana catalog. McGrath and Durand used their analysis for collection evaluation and development. Other authors expanded the techniques, applications, and sources of course subject assignment, which is referred to variously as: (1) course information analysis (Sayles 1984), (2) course analysis (Leighton 1995), or (3) course and research analysis (Lochstet 1997). More recently, authors have studied syllabi, without assigning subject headings, for collection evaluation and collection development (Shirkey 2011).

Syllabi studies provide data on inclusion of library resources, current library use, and to anticipate future library use (Rambler 1982; Lauer, Merz, and Craig 1989; Dewald 2003). Syllabi are studied to identify opportunities for outreach and instruction, particularly in the area of information literacy (Van Scoy

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

and Oakleaf 2008; Morris, Mykytiuk, and Weiner 2014). Some librarians investigate ways to meet cultural heritage curricular needs by conducting systematic although not automated analysis of course catalogs (McGrath and Durand 1969; Leighton 1995; Lochstet 1997; Pancheshnikov 2003). Librarians plan services, resource acquisition using faculty publications (Wormell 2003; Nicholson 2006) and syllabi and other learning objects (Rambler 1982; Sayles 1984; Anderson 1988; McDonald and Micikas 1990; Bean and Klekowski 1993; Shirkey 2001; Dewald 2003), particularly in the area of information literacy (Williams et al. 2004; Van Scoy and Oakleaf 2008; Smith et al. 2012; Hubbard and Lotts 2013; Boss et al. 2014; Morris et al. 2014). Charles (2015) advocates collaboration with educators, librarians, and administrators using "curriculum mapping" to match curricular objectives to assessment in information literacy instruction. Lacking in the literature of either realm, cultural heritage and education, is the discussion of an automated system to link the two.

Case study

In the spring of 2014, Sonia Yaco, head of Special Collections and University Archives at UIC, began exploring more ways to connect academic programs with relevant collections as part of a collection assessment project she was designing. Adding newly designated collecting areas to individual collection's metadata would help faculty get a general sense of appropriate collections. However, how useful is a list of over one hundred political collections to faculty or students? Yaco realized that looking for political science courses that could appropriately use these collections also is problematic. Which of the dozens of courses would be a match for any collection? Yaco began envisioning a new discovery paradigm that could map metadata for collections to curricular metadata. Such a paradigm would be of use to libraries and archives, but other cultural heritage repositories as well. Referred to as the Cultural Heritage and Curriculum Crosswalk, the crosswalk would gather curricular and collections data to match cultural heritage collections to courses. Faculty and students could search on a course title and find relevant material. Students who were interested in a particular collection could search for related courses. Repository administrators, wanting to demonstrate the utility of their holdings to a specific college, could utilize the crosswalk. Development professionals could use the crosswalk to show donors how collections are marketed to patrons. Archivists would use it to enhance existing outreach efforts.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

As proof of concept for the crosswalk, Yaco conducted a pilot study of 101 social justice-related courses taught in a cross-section of UIC colleges and departments. Yaco chose social justice courses because UIC's Special Collections and University Archives Department has strong holdings in social reform, ranging from the Jane Addams Hull House papers to the Chicago Urban League.

The pilot had three goals:

- to collect data for pilot courses and special collection holdings
- to assign authority-controlled subjects to courses
- to find relevant manuscript collections for courses

Methodology

At UIC, gathering course data begins with identifying and gaining access to multiple sources. The online UIC Catalog and Course Request System (CRS) supplied narrative descriptions for courses. Assigned physical and digital course material, which we will refer to generically as "textbooks," came from a variety of sources. The Schedule of Classes lists textbooks for some courses; the Course Reserves in the University Library catalog and the UIC Bookstore list textbooks for courses as well. Each of these sources had a different file format and required a technique to extract the data. Data managers for some of the sources, such as CRS, were able to export data and provide Excel files and Access tables. For other sources, Yaco cut and pasted or retyped data. She cleaned and merged the data using Excel and Microsoft Visual Basic for Applications. Yaco created a Microsoft Access database and imported the gathered data, including a list of the 101 pilot courses, narrative data for one hundred records and textbook data for fifty-two courses. The fifty-two courses have 281 total and 264 unique textbooks.

Yaco used two methods to assign authority-controlled subject terms to courses: harvesting textbook subjects and asking a UIC cataloging librarian to evaluate course descriptions. Yaco hypothesized that the subjects of assigned textbooks would be a good indicator of course content. To compile subjects for these titles, Yaco searched internal and external library catalogs and commercial databases to find the correct titles, International Standard Book Numbers (ISBNs), and subjects; cut, pasted, parsed and cleaned the data with Excel and OpenRefine; imported the subjects into Microsoft Access; and finally connected subjects to titles to courses in the. While this process could have been streamlined by using batch subject lookups in the Online Computer Library Center's (OCLC) Connexion software, either method requires multiple steps of data

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

cleaning. Because of the labor-intensive nature of the lookups, Yaco chose not to obtain subjects for all textbooks assigned to pilot courses. She looked up subjects for 140 of the 264 textbooks, netting 618 unique subjects for fifty-two courses. The librarian assigned subject terms for four courses, adding another eight subjects. In total, there were 626 unique subjects for fifty-five courses (one course had subjects assigned by the cataloging librarian and derived from textbooks).

Manuscript collections are described at UIC in two separate systems. The Special Collections and University Archives department uses Archivists' Toolkit software to create and generate Encoded Archival Description (EAD) finding aids for manuscript collections. Staff comprised of archivists, clerical employees, graduate assistants, and undergraduate students, assigns subjects to collections as part of processing. Subject terms are primarily from Library of Congress Subject Headings (LCSH) but also from the Thesaurus for Graphic Materials, Medical Subject Headings (MeSH), and a locally defined list. Collections with multiple or large finding aids may have multiple resource records. Archivists' Toolkit at UIC has 614 resource records described by 1,627 subjects.

The University Library catalog, a Voyager system, provides intellectual access to general holdings and manuscript collections. Librarians and clerical employees assign LCSH terms to manuscript collections and other holdings. The catalog includes MAchine-Readable Catalog (MARC) records for 522 manuscript collections with 1,521 subject terms. The MARC records include collection-level narrative descriptions from the scope and contents and biography/history portions of finding aids. Because of the differences in the catalogs, Yaco used manuscript metadata from both catalogs for the pilot.

Yaco used a series of queries in Microsoft Access to match the identified course-subjects with collection-subjects in Archivists' Toolkit and collection-subjects and descriptions in the library catalog. To facilitate queries, she normalized subject terms by removing punctuation. After she generated the matches, she removed duplicate recommendations caused by a collection and a course having more than one subject in common. For instance, the collection "Gary Urban League Records" was recommended twice for the course titled "Social Work in a Multicultural Society."

Findings and results

Yaco was able to collect data for all courses, but some courses lacked textbooks or narrative data, which limited the ability to assign subject terms and ulti-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

mately to match their content to collections. Using textbook subjects and direct cataloging of courses, Yaco created authority-controlled subjects for fifty-five of the 101 pilot courses. By querying course-subjects against library catalog terms, collection abstract and scope/contents narrative fields, and Archivists' Toolkit subject terms, Yaco created 885 recommendations for almost a third of the pilot courses (29/101).

Many of the recommendations seemed appropriate. To test library staff's view of relevancy, UIC archivist Gretchen Neidhardt evaluated the recommendations for humanities and social science courses and found about seventy percent (385/544) to be relevant. Yaco and Neidhardt analyzed the recommendations that Neidhardt rated as irrelevant. A few were false hits. The computer matched a course subject *race* to a collection description containing the word *grace*. Other recommendations were inappropriate because the subjects that matched were too broad, such as geographic headings, *women*, and *Americans*. In only one case did the textbook derived subject lead to a recommendation that Neidhardt judged to be irrelevant, a collection of post-modern theater programs for a course on urbanization and gentrification. Neidhardt noted that recommendations would be more useful with a relevancy ranking/rating along with the type of match such as (1) subject term, description, or both; (2) the matching subject term(s); and, (3) the number of matched terms.

Yaco created a report of recommendations by course, listing the matched subject term. She sent emails with collection recommendations to instructional archivists and subject bibliographers. Liaisons and instructional staff have begun using these recommendations in outreach and instruction.

From the pilot to a broader crosswalk

The pilot demonstrates one way to associate collections and curriculum, using textbook subjects to assign subject terms to courses and creating simple Microsoft Access queries to match course-subjects to collection-subjects. Thus far only librarians and archivists have used the recommended collections. However, the crosswalk could be expanded in a variety of ways. Collections from other types of cultural heritage institutions could be linked to courses. A number of stakeholders could benefit from a software implementation of the crosswalk. Faculty planning research classes could use recommendations of relevant primary sources to include in their syllabi and their students could use recommendations when writing papers. Development staff could use the software to show donors of manuscript collections which courses might use their donation. Simi-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

larly, library directors or university administrators could use the software to demonstrate the link between their holdings and programs of study. Knowing the collections related to a course could help academic advisors or student understand chose classes. Moreover, linking metadata from the collections and curricular domains would have a combinatory effect on our knowledge of the content of each.

To develop crosswalk software that could serve this variety of purposes we need to provide specifications and goals using a structure that is understandable by all stakeholders. An ontology, for example, can be used to conceptualize this integrated problem domain (Gruber 2008). This is because an ontology can systematize the description of a complex problem domain (Cimino 2006), such as the one obtained by integrating the curriculum and collections domains. The challenge is to construct an ontology that is a logical, parsimonious, and a complete description of the domain. "We adopt, at least insofar as we are reasonable, the simplest conceptual scheme into which the disordered fragments of raw experience can be fitted and arranged" (Quine 1961, 16). Ramaprasad and Syn (2014a and 2014b) developed a method for breaking a problem domain into logical parts. Using this technique, we developed an ontology of Cultural Heritage and Curriculum Crosswalk shown in Figure 1. (See Appendix A for the ontology glossary).

Fig. 12.1: Curriculum Heritage and Curriculum Crosswalk: An Ontology

The Crosswalk ontology has six main dimensions, which are (1) functions of the system, (2) semiotics of the system, (3) matched cultural heritage, (4) matched curriculum, (5) agents who need/use the system, and (6) outcomes of the system. The underlying argument is that the system through its functions and semiotics will help the agents match the cultural heritage and the curriculum to achieve the desired outcomes.

The Functions of the system required for crosswalk include the traditional information systems functions for managing the semiotics to match the cultural

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

heritage and the curriculum. The Functions are (1) Acquire/collect, (2) Create/edit/delete, (3) Organize/arrange, (4) Index/describe/represent, (5) Store/preserve/ sustain, (6) Secure/authenticate, (7) Retrieve, (8) Process, and (9) Distribute.

Semiotics of the Crosswalk include (1) Data, (2) Information, and (3) Knowledge about the cultural heritage and the curriculum. "Semiotics describes the process by which stimuli are transformed into information and information is transformed into stimuli; in other words, it is the process by which information is generated and dissipated" (Ramaprasad and Rai 1996, 187). The Data include measurements and observations, both qualitative and quantitative, about Cultural Heritage and Courses, such as library catalog records and course descriptions. The Information includes relationships among the data about, and between, Collections of Cultural Heritage and the Curriculum. The Knowledge includes Pedagogical and Cultural interpretation of the above relationships with reference to each and matching the two. The Semiotics elements are ordinal: Information is derived from Data, and Knowledge is derived from Information.

Cultural heritage is broadly classified as "Tangible" and "Intangible" by UNESCO (1970, 2003). The Tangible artifacts include physical Archeological, Prehistorical, Historical, Literary, Artistic, and Scientific objects. The majority of holdings at libraries and archives are Tangible-Literary, "rare manuscripts and incunabula, old books, documents and publications of special interest (historical, artistic, scientific, literary, etc.) singly or in collections" (UNESCO 1970, 4). Museum holdings could be any of the Tangible types. The Intangible cultural heritage includes "(a) oral traditions; (b) performing arts; (c) social practices, rituals, and festive events; (d) knowledge and practices concerning nature and universe; and traditional craftsmanship" (UNESCO 2003, "Introduction").

The Curriculum of the system includes the (1) Syllabi for the courses, (2) course Content, (3) course Textbooks, (4) planned Learning Outcomes, (5) the Faculty members teaching the courses, (6) the Students in the courses, (7) the Pedagogy used, and (8) the Assessment of the courses. A curriculum on a topic may have these elements in different measures with different emphases. Not explicitly listed in the curriculum, but contained in some courses, are university-wide curricular goals and mandates such as undergraduate learning objectives. The elements provide the content for matching curriculum with the cultural heritage. The elements are not listed in any particular order.

A variety of Agents play a role in matching cultural heritage and the curriculum to obtain the desired outcomes. Cultural heritage professionals and Information systems professionals include librarians, archivists, curators, and digital services staff at repositories. Other Agents are (1) Students, (2) Faculty members, (3) Academic advisors, (4) Administrators (Cultural heritage institution

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

and University), (5) Financial stakeholders (such as Legislators, Development/advancement staff, and Donors), and (6) the General Public. Any of these Agents may contain alumni.

Lastly, the potential Outcomes of the crosswalk are many. These may range from investigating (Investigate) to enjoying (Enjoy) cultural heritage. They include (1) Investigate, (2) Document, (3) Preserve, (4) Manage, (5) Visualize, (6) Educate, (7) Communicate, (8) Access, (9) Assess, (10) Plan, (11) Publicize, (12) Contextualize, and (13) Enjoy.

The proposed ontology is a lens to study the anatomy of the crosswalk. For a complex problem like matching two domains, there may be other lenses to study the problem with each encapsulated by a different ontology. They will provide different perspectives. We will discuss the present one in detail.

We have discussed the individual dimensions (columns) and elements (rows) of the ontology while describing the construction of the ontology. Multiple elements of a dimension may coexist independently but may also interact with each other in the Crosswalk. Thus, Functions, Semiotics, Cultural heritage, Curriculum, Agents, and Outcomes may coexist and interact with each other. Knowing the independent and interacting elements is critical to envisioning and implementing crosswalk.

The dimensions of the ontology are arranged left to right with adjacent words/connectors such that the concatenation of an element from each dimension with the adjacent words/connectors creates a natural English sentence illustrating a potential component of the Crosswalk. The components define the Crosswalk. The pilot Yaco conducted, for instance, contains these components:

- a system to acquire/collect data-courses to match tangible-literary collections and curriculum textbooks to aid cultural heritage professionals to educate [about] cultural heritage
- a system to index/describe/represent data-courses to match tangible-literary collections and curriculum syllabi to aid faculty members to educate [about] cultural heritage
- a system to retrieve information- collections to match tangible-artistic collections and curriculum syllabi to aid faculty to access cultural heritage

The ontology encapsulates 679,536 potential components of the crosswalk. It would be laborious and voluminous to enumerate all of them. The ontology provides a convenient and concise “big picture” of the crosswalk in a limited space. It helps visualize the combinatorial complexity of the matching problem.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

Practical solutions

Arguably, the success of developing a software solution for the crosswalk is dependent on a clear definition of requirements. The crosswalk ontology provides a comprehensive framework and a common language for stakeholders. The approaches discussed below apply the lessons learned from the pilot and utilize the ontology to guide development or selection of software that associate courses with collections and vice versa for better utilization of cultural heritage resources. All the approaches will need to address local practices for data stewardship, data governance, access permissions, and so forth.

Approach one: Improve on methods and data used in the case study to develop a solution that is specific to UIC and then extend the solution using the ontology to accommodate needs for a generic solution. Working with campus and consortia data administrators would allow access to additional data sets and data extraction tools. Current methods, including data cleaning, could be streamlined to increase the data currency and quality. For instance, at UIC the library and course content managers in the Office of Academic Program Development are discussing having faculty or program directors choose subjects for courses from a short librarian-supplied list. Alternately, the addition of new data sets might reduce the need for authority-controlled subjects in both domains. Using more advanced matching techniques with additional data could increase the number and quality of the search results. One of the major flaws with this approach is that it requires use of its own copy of data sets, which would often be out-of-sync with the live data.

Approach two: Create an application using data mining feature of software like the Microsoft SQL Server. These systems provide tools for data cleaning and preparation, machine learning and reporting all in one platform. Among the advantages are streamlined data processing time and prediction capabilities. Like the previous approach, this would require a huge institutional commitment to maintain a parallel system. Unless efficient and streamlined methods are established to update the data snapshots, this would run the risk of always be out of date.

Approach three: Develop an online tool with an embedded Google search engine pointed at curricular and collections data sets. As part of installing the software, each university would identify the location and authentication protocol of local or consortium data sets. Google's proprietary search algorithms would be used to match each domain's authority controlled subjects. This approach does not require redundant data sets but does need authority controlled subjects to be created for courses.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

Approach four: Create or adapt a recommender system, which uses artificial intelligence algorithms. This approach, used by Amazon and other retailers, would utilize machine-learning algorithms to predict which collections are relevant to which courses. Set up would include specifying the location and authentication protocol of local or consortium data sets. Such a system would not require redundant data sets or structured authority controlled data. Recommender systems are flexible enough to use with any size, format, or number of data sets. The ontology would be particularly useful for evaluating commercial recommender software for this approach.

A parallel use of the crosswalk ontology is mapping the 'bright' (heavily emphasized), 'light' (lightly emphasized), and 'blind/blank' (overlooked unconsciously and consciously) spots in the collection and curriculum to reveal the strengths and weaknesses of the collections and the differential emphases of the curriculum. This mapping can also help to articulate the match and mismatch between the collections and the curriculum. Using these insights, cultural heritage professionals and faculty members can modify the portfolio of collections, modify the curriculum, and optimize the match between the collections and curriculum.

Conclusion

Our rich storehouses of collection metadata, representing decades of work, make access to collections possible from across the world. Yet they are rarely used to reach across campus. The Cultural Heritage and Curriculum Crosswalk presents a new discovery paradigm. The case study at UIC demonstrates one approach to the crosswalk, using metadata to generate automated recommendations of collections for courses. A complicated data environment for courses, collections, and subjects at UIC make the pilot methods of data gathering unscalable. Whether other universities have intricate or straightforward data structures for tracking course content and cultural heritage collections, it is likely that each school has its own unique mix of data sources. The varied environments of holdings and curricular data call for a structure that is designed to handle complex problems. The crosswalk ontology can be used to develop a roadmap for integrating the cultural heritage and curricular domains. In the absence of such a framework, efforts at integration tend to be fragmented and episodic. The former is likely to be more effective than the latter. The approaches suggested in this chapter can build on the ontology to advance the development

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

of curriculum and collections for the benefit of university and cultural heritage stakeholders.

Appendix A

Ontology Glossary

Function: The functions of the system.

- Acquire/collect: To acquire/collect data/information/knowledge about cultural and curriculum.
- Create/edit/delete: To create/edit/delete data/information/knowledge about cultural and curriculum.
- Organize/arrange: To organize/arrange data/information/knowledge about cultural heritage and curriculum.
- Index/describe: To index/describe data/information/knowledge about cultural heritage and curriculum.
- Store/preserve/sustain: To store/preserve/sustain data/information/knowledge about cultural heritage and curriculum.
- Secure/authenticate: To secure/authenticate data/information/knowledge about cultural heritage and curriculum.
- Retrieve: To retrieve data/information/knowledge about cultural heritage and curriculum.
- Process: To process data/information/knowledge about cultural heritage and curriculum.
- Distribute: To distribute data/information/knowledge about cultural heritage and curriculum.

Semiotics: Symbolic representations of cultural heritage and curriculum at different levels of abstraction.

- Data: Measurements and observations, qualitative and quantitative about cultural heritage and curriculum.
  - Cultural heritage: Data about Cultural heritage.
  - Courses: Data about courses in the curriculum.
- Information: Relationships among the data about cultural heritage and curriculum.
  - Collections: Information about collections of cultural heritage.
  - Curriculum: Information about the curriculum.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

- Knowledge: Interpretations of the relationships among the cultural heritage and curriculum data.
  - Cultural: Knowledge about culture.
  - Pedagogical: Knowledge about the pedagogy of the curriculum.

Cultural heritage: Artifacts and heritage identified with a culture.
- Tangible: Tangible cultural heritage artifacts.
- Archaeological: "Products of archaeological excavations (including regular and clandestine) or of archaeological discoveries.â€ (UNESCO 1970)
- Prehistorical: Property predating history such as prehistorical tools, structures, paintings, etc.
- Historical: "Property relating to history, including the history of science and technology and military and social history, to the life of national leaders, thinkers, scientists and artist and to events of national importance.â€ (UNESCO 1970)
- Literary: "Rare manuscripts and incunabula, old books, documents and publications of special interest (historical, artistic, scientific, literary, etc.) singly or in collections.â€ (UNESCO 1970)
- Artistic: "Pictures, paintings, and drawings produced entirely by hand on any support and in any material (excluding industrial designs and manufactured articles decorated by hand); original works of statuary art and sculpture in any material; original engravings, prints and lithographs; original artistic assemblages and montages in any material." (UNESCO 1970)
- Scientific: Scientific and technological objects, instruments, devices, etc. significant to the advancement of science and technology.
- Intangible: Intangible cultural heritage.
- Oral: Oral traditions and expressions, including language as a vehicle of the intangible cultural heritage.
- Performance: Performing arts.
- Social practices/nature: Social practices, rituals and festive events.
- Knowledge/practice nature: Knowledge and practices concerning nature and the universe.
- Traditional craftsmanship: Traditional craftsmanship.

Curriculum: A systematic plan for knowledge transmission, generally in a university.
- Syllabi: A description of the structure of the courses in the curriculum.
- Content: The content of the course.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

- Textbooks: Physical or digital media (including books, audio-videos, articles, maps) embodying the common body of knowledge regarding the subject of the courses in the curriculum.
- Learning Outcomes: The learning outcomes desired from the courses in the curriculum.
- Faculty: The faculty members teaching the courses in the curriculum.
- Students: The students taking the courses in the curriculum.
- Pedagogy: The method of teaching the courses in the curriculum.
- Assessment: The assessment of the teaching of and learning from the courses in the curriculum.

Agents: The agents to obtain the various outcomes regarding cultural heritage.
- Students: The students of cultural heritage.
- Faculty members: The faculty members of cultural heritage.
- Academic advisors: The academic advisors to students and faculty members.
- Administrators: The administrators of cultural heritage outcomes.
- Cultural heritage institution: The administrators of cultural heritage institutions (library, archives, and museums) and the outcomes.
- University: The university administrators of cultural heritage and the outcomes.
- Financial stakeholders: the financial stakeholders in the cultural heritage and outcomes.
- Legislators: The legislators involved in funding the cultural heritage and outcomes.
- Dev. /Adv. Staff: The development and advancement staff members raising funds for the cultural heritage and outcomes.
- Donors: The donors to programs in cultural heritage and outcomes.
- Cultural heritage professionals: the cultural heritage professionals responsible for the outcomes.
- Information system professionals: The information systems professionals responsible for the system and its outcomes.
- General public: Public at large.

Outcome: The outcomes of integrating cultural heritage into the curriculum.
- Investigate: To investigate cultural heritage.
- Document: To document cultural heritage.
- Preserve: To preserve cultural heritage.
- Manage: To manage cultural heritage.
- Visualize: To visualize cultural heritage.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

- Educate: To educate cultural heritage.
- Communicate: To communicate cultural heritage.
- Access: To access cultural heritage.
- Assess: To assess cultural heritage.
- Plan: To plan cultural heritage.
- Publicize: To publicize cultural heritage.
- Contextualize: To contextualize cultural heritage.
- Enjoy: To enjoy cultural heritage.

References

Allison, Anna E. 2005. "Connecting Undergraduates with Primary Sources: A Study of Undergraduate Instruction in Archives, Manuscripts, and Special Collections." Master's Paper for the M.S. in L.S. Degree. https://cdr.lib.unc.edu/.../uuid:cc3864a9–10e4–467e-8b56–8171e400ab7b.

Anderson, Renee Nesbitt. 1988. "Using the Syllabus in Collection Development." *Technicalities* 8: 14–15. http://thescholarship.ecu.edu/bitstream/handle/10342/3613/Using%20Syllabi%20for%20Collection%20Development%20User-Centered%20Collection%20Development.doc?sequence=1.

Bean, Rick Jon and Lynn Marie Klekowski. 1993. "Course Syllabi: Extracting their Hidden Potential at DePaul University's Suburban Campus Libraries." The Off-Campus Library Services Proceedings, Central Michigan University Press; Mt. Clemens, Michigan.

Boss, Katherine and Emily Drabinski. 2014. "Evidence-Based Instruction Integration: A Syllabus Analysis Project." *Reference Services Review* 42 (2): 263–276. doi:10.1108/RSR-07–2013–0038.

Charles, Leslin H. 2015. "Using an Information Literacy Curriculum Map as a Means of Communication and Accountability for Stakeholders in Higher Education." *Journal of Information Literacy* 9 (1): 47–61. doi:10.11645/9.1.1959.

Cimino, James J. 2006. "In defense of the Desiderata." *Journal of Biomedical Informatics* 39, 299–306. doi:10.1016/j.jbi.2005.11.008.

Dewald, Nancy H. 2003. "Anticipating Library use by Business Students: The Uses of a Syllabus Study." *Research Strategies* 19 (1): 33–45. doi:10.1016/j.resstr.2003.09.003.

Gruber, Thomas R. 2008. "Ontology." In *Encyclopedia of Database Systems*, L. Liu and M.T. Özsu, eds. Springer-Verlag.

Hubbard, Melissa A. and Megan Lotts. 2013. "Special Collections, Primary Resources, and Information Literacy Pedagogy." *Communications in Information Literacy* 7 (1): 24–38. doi:10.7282/T3028PFX.

Kennedy, Louise. 2006. "Partners or Gatekeepers? Archival Interactions with Higher Education at University College Dublin." In *Archives and Archivists*, edited by Ailsa C. Holland and Elizabeth Mullins, 219–231. Dublin Ireland: Four Courts Press.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

Kesselman, Martin A., and Sarah Barbara Watstein. 2009. "Creating Opportunities: Embedded Librarians." *Journal of Library Administration* 49 (4): 383–400. doi:10.1080/01930820902832538.

Kvenild, Cassandra; Calkins, Kaijsa. 2011. *Embedded Librarians: Moving Beyond One-Shot Instruction*. Chicago: Association of College and Research Libraries.

Lauer, Jonathan D. 1989. "What Syllabi Reveal About Library Use: A Comparative Look at Two Private Academic Institutions." *Research Strategies* 7 (4): 167–174.

Leighton, H. Vernon. 1995. "Course Analysis: Techniques and Guidelines." *Journal of Academic Librarianship* 21 (3): 175–179.

Lochstet, Gwenn S. 1997. "Course and Research Analysis using a Coded Classification System." *The Journal of Academic Librarianship* 23 (5): 380–389. doi:10.1016/S0099–1333(97)90035–5.

Maher, William J. 2001. *The Management of College and University Archives*. Metuchen, N.J.: Scarecrow Press.

McDonald, Joseph A. and Lynda Basney Micikas. 1990. *Acquisitions '90: Conference on Acquisitions, Budgets, and Collections*, May 16 and 17. St. Louis, Missouri: Proceedings. United States.

McGrath, William E. and Norma Durand. 1969. "Classifying Courses in the University Catalog." *College & Research Libraries* 30 (6): 533–539. doi:10.5860/crl_30_06_533.

Montoya, Robert D. 2013. "Teaching and Learning Los Angeles through Engagement with UCLA Library Special Collections." *Urban Library Journal* 19 (1): 1–12. http://academicworks.cuny.edu/ulj/vol19/iss1/3/.

Morris, Sammie, Lawrence Mykytiuk, and Sharon Weiner. 2014. "Archival Literacy for History Students: Identifying Faculty Expectations of Archival Research Skills." *American Archivist* 77 (2): 394–424. doi:10.17723/aarc.77.2.j270637g8q11p460.

Nicholson, Scott. 2006. "The Basis for Bibliomining: Frameworks for Bringing Together Usage-Based Data Mining and Bibliometrics through Data Warehousing in Digital Library Services." *Information Processing & Management* 42 (3): 785–804. doi:10.1016/j.ipm.2005.05.008.

Pancheshnikov, Yelena. 2003. "Course-Centered Approach to Evaluating University Library Collections for Instructional Program Reviews." *Collection Building* 22 (4): 177–185. doi:10.1108/01604950310501762.

Quine, Willard Van Orman. 1961. *From a Logical Point of View*. Boston: Harvard University Press.

Ramaprasad, A., and A Rai. 1996. "Envisioning Management of Information." *Omega – International Journal of Management Science* 24 (2): 179–79. doi:10.1016/0305–0483(95)00061–5.

Ramaprasad, Arkalgud and Thant Syn. 2014a. "Design Thinking and Evaluation Using an Ontology." In *Design Science: Perspectives from Europe*. Edited by M. Helfert, B. Donnellan and J. Kenneally Eds. Springer International Publishing: 63–74.

Ramaprasad, Arkalgud and Thant Syn. 2014b. "Ontological Topography: Mapping the Bright, Light, Blind/Blank Spots in Healthcare Knowledge." In *Proceedings of Big Data Analysis in Healthcare* 2014, Singapore. https://www.researchgate.net/publication/281493992_Ontological_Topography_Mapping_the_Bright_Light_BlindBlank_Spots_in_Healthcare_Knowledge.

Rambler, Linda K. 1982. "Syllabus Study: Key to a Responsive Academic Library." *The Journal of Academic Librarianship* 8 (3): 155–159.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

Robyns, Marcus C. 2001. "The Archivist as Educator: Integrating Critical Thinking Skills into Historical Research Methods Instruction." *American Archivist* 64 (2): 363–384. doi:10.17723/aarc.64.2.q4742x2324j10457.

Rockenbach, Barbara. 2011. "Archives, Undergraduates, and Inquiry-Based Learning: Case Studies from Yale University Library." *American Archivist* 74 (1): 297–311. doi:10.17723/aarc.74.1.mml4871x2365j265.

Sayles, Jeremy W. 1985. "Course Information Analysis: Foundation for Creative Library Support." *The Journal of Academic Librarianship* 10 (6): 343.

Shirkey, Cindy. 2011. "Taking the Guesswork Out of Collection Development: Using Syllabi for a User-Centered Collection Development Method." *Collection Management* 36 (3): 154–164. doi:10.1080/01462679.2011.580046.

Smith, Cheri, Linda Doversberger, Sherri Jones, Parker Ladwig, Jennifer Parker, and Barbara Pietraszewski. 2012. "Using Course Syllabi to Uncover Opportunities for Curriculum-Integrated Instruction." *Reference and User Services Quarterly* 51 (3): 263–271. doi:10.5860/rusq.51n3.263.

UNESCO. 1970. *Convention on the Means of Prohibiting and Preventing the Illicit Import, Export and Transfer of Ownership of Cultural Property 1970*. http://portal.unesco.org/en/ev.php-URL_ID=13039&URL_DO=DO_TOPIC&URL_-SECTION=201.html.

UNESCO. 2003. *Article 2 – Definitions. Intangible Cultural Heritage*. www.unesco.org/culture/ich/en/convention#art2.

VanScoy, Amy and Megan J. Oakleaf. 2008. "Evidence vs. Anecdote: Using Syllabi to Plan Curriculum-Integrated Information Literacy Instruction." *College and Research Libraries* 69 (6): 566–575. doi:10.5860/crl.69.6.566.

Weiner, Sharon Gray. 2005. "The History of Academic Libraries in the United States: A Review of the Literature." *Library Philosophy and* Practice 7 (2). http://digitalcommons.unl.edu/libphilprac/58.

Williams, Lisa M., Sue Ann Cody, and Jerry Parnell. 2004. "Prospecting for New Collaborations: Mining Syllabi for Library Service Opportunities." *The Journal of Academic Librarianship* 30 (4): 270–275. doi:10.1016/j.acalib.2004.04.009.

Wormell, Irene. 2003. "Matching Subject Portals with the Research Environment." *Information Technology and Libraries* 22 (4): 158–164.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

K. Meghan Gross, Cory Lampert, and Emily Lapworth

13 Optimizing Merged Metadata Standards for Online Community History: A Linked Open Data Approach

Introduction

Large-scale digital library projects began emerging, first as research experiments in the mid-1990s, and later as federally funded initiatives that sought to model the services and community outreach practices of traditional libraries (Besser 2004). Partnerships like the Digital Library Federation, formed in 1995, initiated early efforts on open standards and metadata interoperability and over the years digital collections have flourished across cultural heritage institutions, including in academic libraries with rare unique materials. Nearly every sector of the information professions has a mandate to manage digital assets and their associated metadata. Merging metadata of varying quality and from diverse repository systems is a necessary and essential function of digital collections work. Data is an asset with one possible application in the field of linked open data (LOD). By transforming metadata into semantic web-based standard data formats and assigning Uniform Resource Identifiers (URI), rich metadata describing digital collections can reveal contextual relationships and be utilized by computer systems without human intervention. The advent of LOD in libraries has revealed the value of library data and influenced perspectives on metadata quality. In this chapter the authors will examine the complexities of merging metadata from myriad sources with a focus on metadata management practices that add value while balancing the need for efficiency and collaborative workflow.

Since 2006, University of Nevada, Las Vegas (UNLV) Libraries have been successful in securing external funding to build comprehensive topic-focused digital collections that document the history of Southern Nevada. These online portals contain digitized objects, standards-compliant metadata from multiple sources, and contextual narrative delivered via a digital asset management system that allows users to access and interact with the content. Metadata management for these projects has been subject to a wide range of variables including (1) legacy practices, (2) staffing resources, (3) multiple systems and technologies, and (4) evolving user expectations. As the digitization program has followed national trends toward more rapid production, workflow approaches

https://doi.org/10.1515/9783110337419-014

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

have evolved to meet growing needs and maximize production quantity. Using the Southern Nevada Jewish Heritage Project (SNJHP) as a case study, this chapter will explore how three key areas of metadata management (1) legacy practice, (2) implementation of library systems, and (3) linked open data (LOD), have precipitated collaboration, negotiation, and philosophical discussion about the need for metadata strategies that address sustainability while allowing for innovation in metadata creation to add value for users.

Legacy practice and systems for managing metadata

The source of many UNLV Digital Collections priorities is the Special Collections division of the Libraries, which is home to a diverse set of materials (manuscript collections, university archives, photograph collections, oral histories, etc.) documenting the history of gaming and Southern Nevada. An active digitization program at UNLV Libraries was developed in 2006 to enhance user access and expose rare primary resources to a global audience. The program was initiated and sustained without a centralized, searchable content management system (CMS) or dedicated technical services department serving the Special Collections division. The primary method of obtaining metadata for digitization projects was by recording extant metadata from paper inventories or through the human intervention of a curator who would select material manually by topic and research value, and prioritize it for digitization based on themes. Often the curator would answer basic questions about "collection names," "creator," or "related items" as requested by the metadata librarian in the Digital Collections department in order to assign required values in the Dublin Core (DC) metadata created for online access.

In recent years, UNLV Special Collections has undertaken two efforts that have fundamentally shifted metadata management practices. A reorganization took place that created two departments in Special Collections: (1) Technical Services and (2) Public Services, and moved a third, (3) Digital Collections, from the Library Technologies division to the Special Collections division. Now fully staffed, the departments have potential to rethink workflows and work in tandem to understand metadata goals for both archival researchers and the global web audience. Though these goals may at times focus on different priorities, the commitment to increased access is shared across the division.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

Secondly, the UNLV Special Collections Technical Services department implemented the ArchivesSpace collections management software shortly before the start of the SNJHP. Of particular interest to metadata managers is the way that ArchivesSpace provides a unified content management system for oral histories, manuscript collections, photograph collections, and university archives. Within one system, finding aids are linked to accession records and agent records, creating connections that are easily visualized and searchable within the system. Staff across the division, not just within the Technical Services department, can easily access and contribute to the information in ArchivesSpace. Division-wide access alleviates the challenge of establishing priorities defining which materials deserve enhanced description and which should be digitized with only finding aid metadata. There are also promising benefits that result from the adoption of recognized archival metadata standards. The ArchivesSpace interface helps staff input description that conforms to Describing Archives: A Content Standard (DACS) and can be exported in Encoded Archival Description (EAD) for reuse.

One potential reuse is in the Digital Collections digital library system, CONTENTdm. This content management system provides a back-end cataloging interface, as well as a public web interface. To date (2016), UNLV Libraries' CONTENTdm installation houses twenty-one distinct digital collections made up of over 100,000 items. Item types range from newspapers to "born-digital" photographs, and each of the item types presents an opportunity to provide metadata in a way that can be retrieved through simple or advanced searches. CONTENTdm supports a relatively flat data structure for describing and displaying digital objects and this is often in conflict to hierarchical metadata, such as the way archival collections are described. Hyperlinks are made available within the description of digital objects that allow a user to browse beyond a defined list of search results, but the relationship between and among records is not specified (beyond a "source" field). Controlled vocabularies (CV) are a key component of CONTENTdm. In designated fields data entry is restricted to these pre-defined (CV) terms, which aids in assigning subjects, retrieving precision search results, sharing metadata between systems, and ultimately, for transforming metadata into linked data.

Because the majority of prior digitization projects were undertaken with minimal access to standards-based archival metadata, the Digital Collections department developed a parallel process to archival methods that prepared digital objects for enhanced access through digitization. Rich, granular, and primarily item-level metadata was created in alignment with national best practices for digital libraries using the Dublin Core metadata standard and a locally

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

created UNLV Application Profile, which ensured compliance and rigor in metadata application, both locally and regionally.

In contrast to the well-established item-level curatorial approach, UNLV Digital Collections also piloted full-collection digitization and display through the CONTENTdm interface. With methods based on mass-digitization projects at other institutions such as Archives of American Art (Aikens et al. 2011) and North Carolina State University (Dietz and Ronallo 2011), the goal was to programmatically address discovery issues of special collections materials. In these projects, the information captured in the finding aid, usually in Encoded Archival Description (EAD), is used to describe the digitized items; therefore, item-level description can be reduced or eliminated. At UNLV, differing standards and levels of description, in addition to the digital library system in place (CONTENTdm), prevented Digital Collections from digitizing and describing full archival collections using only EAD. The department found that these same problems persist in the SNJHP, in which merging collection level and item-level metadata is a priority.

Merging metadata for the Southern Nevada Jewish Heritage Project

Concurrent to UNLV's discussions about full-collection digitization and the implementation of ArchivesSpace, the Southern Nevada Jewish Heritage Project was initiated in 2014 to collect and provide increased online access to historical materials specifically about the Jewish community in Southern Nevada. The SNJHP web portal, a major project outcome, provides online digitized primary sources alongside related contextual information such as finding aids, biographical information, and a historical timeline. Before the project began, UNLV Special Collections already had about thirty archival collections and fifty oral history interviews related to the local Jewish community. Selections from these existing resources were digitized, and over the next two years project staff acquired twenty new manuscript collections and over one hundred new oral history interviews, which were also digitized and added to the web portal.

As the Technical Services department set out to improve description and access of these collections using archival methods, the Digital Collections staff simultaneously set out to achieve the same goals by digitizing and assigning metadata to materials selected from new and existing manuscript, photograph, and scan-and-return collections. Due to backlogs and deadlines for the project,

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

not all materials for the SNJHP had been prepared with consistent, uniform metadata that could be directly repurposed for digital collections. The ideal flow of metadata through the Special Collections division can be gleaned from Table 13.1.

Tab. 13.1: Metadata Flow through Special Collections Division.

People	Curator	Accessioning archivist	Processing archivist	Digitization staff
Records	Curator's form	Accession record	Finding aid	Digital object metadata
Systems			ArchivesSpace	CONTENTdm
Standards			DACS/EAD	Dublin Core UNLV Application profile
Levels of description	Accession	Accession	Collection (→ file)	Object (→ item)

In addition to the challenge of migrating metadata across systems and standards (from ArchivesSpace/EAD to CONTENTdm/Dublin Core), the consolidation of collection information into ArchivesSpace is not yet complete, meaning that some materials are queued for digitization with little or no standardized metadata. The SNJHP was faced with several pipelines of materials with varying levels of access and a diverse set of existing metadata. The compressed time frame of the SNJHP and the unconventional method of scan-and-return collecting are also challenges to efficiently merging metadata from multiple sources to create an online digital collection and LOD interface.

Formats and levels of description

One of the biggest challenges in legacy metadata, beyond the lack of readily available collection information, is the varying degree of description applied to various formats. The format, or quantities associated with a format, often dictates the pipeline through which the material enters the metadata workflow. Through this workflow a level of description must be selected as well as the key access points to target in metadata creation. Capturing "aboutness," or the totality of subjects explicitly or implicitly addressed in an object, can be quite different based on formats; for instance, the description of the visual content of what a photograph shows (people, places, events, or even situations) versus the publication information of a book with a full text transcript. For photographs,

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

metadata must be supplied; for a full text transcript, the metadata can be extracted with indexing systems.

Visual materials have always been a high priority for user access at UNLV. This is reflected in the fact that UNLV's photograph collections were historically described at the item-level in paper inventories as well as in a homegrown online database. Oral histories were tracked in yet another database, and transcripts, audio, and video assets were managed outside of any central system. As in any repository, backlogs of unprocessed materials wait in hopes of prioritization for any form of access and description. Some of these collections were also queued for digitization, despite the lack of any formal metadata or collection assignment.

Manuscript collections, like those selected for inclusion in the SNJHP, received varying levels of description, from those that were not processed at all, to collection level descriptions in a database, to folder level inventories, to the more recent standard of DACS compliant finding aids, which can range from collection level to file level descriptions. Varying levels of description is a challenge that digital collections staff will continue to face, especially since the profession has largely accepted the strategy proposed by the ground breaking article, "More Product, Less Process: Revamping Traditional Archival Processing." In order to provide basic access and to diminish seemingly insurmountable backlogs, the authors suggest that processing below the collection level should be the exception rather than the rule, and that the level of processing should be determined on a case-by-case basis (Greene and Meissner 2005).

The standards for describing archival collections differ from those for describing digital collections. Archival description uses a hierarchical format, so that the highest level of information can trickle down to the item-level. However, digital collections metadata practice relies on disaggregated item-level description to be searchable/retrievable, and the UNLV Libraries' digital library system, CONTENTdm, has a flat structure that limits hierarchical description. Eliminating item-level description decreases the opportunity for digital collections to bring together related items from disparate donors, manuscript collections or formats. Therefore, UNLV Digital Collections still sees a need to perform some item-level digitization and description.

"Scan-and-return" collections

Another challenge related to metadata creation and interdepartmental workflow for the SNJHP is the method of "scan-and-return" collecting, in which case

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

items are digitized and described by Digital Collections at the item-level before Technical Services accessions their surrogates. In the SNJHP, not all community members are willing to transfer physical custody of their records, but creating surrogates allows Special Collections to provide long-term, online, public access to these important resources. The challenge inherent to this process is that appraisal, digitization, and description are performed at the item-level before the materials are accessioned and described at the aggregate level. Questions arose about which staff performed which tasks, and how to best avoid duplication of effort. Initially, the project archivist appraised items and created a draft finding aid, then the digitization specialist digitized and created item-level metadata that adhered to the UNLV Application Profile, after which the accession information and finding aid could be finalized and published. As project staffing was reduced, the digitization specialist was empowered to make appraisal decisions and notify the division's accessioning archivist when digital surrogates, complete with their item-level metadata, were ready for accessioning and collection level description by the Technical Services department.

Audience

The SNJHP allows users to explore materials at the collection level via finding aids created in ArchivesSpace or at the item-level via the digital collection housed in CONTENTdm. Different systems are designed for different formats and metadata types, but it is not just the metadata workflow that must be taken into account when deciding what to make available, how, and where. The intended audience for the material and those users' patterns of online search behavior also informs this decision. Based on the audience of the SNJHP, and the richness and potential of the existing metadata, UNLV Special Collections decided to repurpose this metadata to create additional access points based on agents and linked open data.

Agent's metadata

The need to highlight and expose connections between people led to deeper investigation into how our systems managed agents (people and groups) through (1) controlled vocabularies, (2) standards-based metadata, and (3) ontologies that described relationships. A stated goal of the SNJHP project was to focus on

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

agents and their connections as a primary access path to the digitized content as well as a navigation tool for the web portal. Metadata work that focused on agents took place in several systems to realize this goal: (1) in a spreadsheet of people identified as potential oral history interviewees, (2) in ArchivesSpace when collections were accessioned and processed, (3) in CONTENTdm as metadata records were created, (4) in TemaTres vocabulary management system for local names, and (5) in a Drupal module developed to capture agent information for reuse on the portal website.

ArchivesSpace staff users can easily see which resources are associated with which agents, which agents are associated with each other, and how. However, the public interface of ArchivesSpace will not be implemented by UNLV until 2017, so SNJHP project staff replicated and enhanced these connections in the project web portal. The web portal includes biographies of notable Jewish people and historical summaries of organizations and congregations, all with links to related finding aids, oral history audio clips, transcripts, and digital objects. When compiling these links, ArchivesSpace made it easy for staff to see all resources related to an agent. Even though ArchivesSpace is not accessible to the public yet, there is still value in the utility of the program because it has made internal workflows more efficient while linking related data elements via the SNJHP web portal. The oral historian can, for example, record biographical information about oral history interviewees that the archivist can reuse in the biographical note of a finding aid. Later, the digitization specialist can refer to the finding aid to enhance the efficiency of creating item-level metadata for a digital object.

In 2013, UNLV Libraries Digital Collections department began investigating the implementation of linked open data with the collections represented in CONTENTdm. Many digital collections conceived of as "projects" began to overlap with each other thematically and referenced the same key individuals and groups. In order to facilitate the visualization of the relationships between and among objects, and the people represented in digital objects such as authors, interviewers, and subjects, the digital collections metadata librarian undertook efforts to audit the data from CV fields in the collections for accuracy. Using Google's OpenRefine, controlled vocabulary terms were normalized, verified, and re-imported into the controlled vocabulary and re-applied to their assets.

The method for capturing names and other controlled vocabulary terms for inclusion in the digital collections at UNLV Libraries evolved in several phases, from direct input into the vocabulary through the CONTENTdm administration module, to sending an email to an administrator-level user suggesting new terms, to capturing terms and their sources in a spreadsheet, which were added to the CONTENTdm controlled vocabulary by an admin-level user. The spread-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

sheet method was helpful in creating an interim local controlled vocabulary, with research notes and known Uniform Resource Identifiers (URIs), but was unwieldy for multiple users. Using this method highlighted several issues: (1) duplicate entries in the CV were not eliminated due to user error while searching for previously entered terms, (2) there was no way to capture or effectively search for alternate names, and (3) there was no way to capture relationships among entries. All of these issues directly affect the implementation of linked data.

To eliminate the problems created and perpetuated through the interim spreadsheet, UNLV Digital Collections adopted an internal open-source controlled vocabulary server, TemaTres. TemaTres is highly customizable as an open source application and the benefits to using the controlled vocabulary server are numerous (TemaTres 2016). The team now has better control over local names, which make up the bulk of the original metadata they create, and can generate stable URLs through the software. Another tangible impact comes in the form of capturing the relationships among terms stored in the database. It is within this server/database that the team can link organizations with people, parents with children, colleagues, etc. These relationships, expressed in linked open data with associated URIs, can be leveraged to provide user experiences searching and browsing information sources free of system data silos and with vastly more contextualized results.

Metadata to linked open data for user consumption

LOD is defined as a set of best practices for publishing and connecting structured data on the web (Bizer, Heath, and Berners-Lee 2009) and some of the most interesting potential of this work is visualizing data and revealing not just that information is related, but how it is related. Capturing the rich connections within and between research content that cultural heritage repositories hold is a natural extension of the traditional work of catalogers and metadata creators. UNLV Digital Collections developed several principles that guide local practice and empower those working most closely with the content, the data contributors, to preserve the richness of their contributed data while leveraging web technologies to transform their metadata into linked open data. Because the SNJHP focused so heavily on community connections between people and groups and was developed in such a close collaboration between curators, oral

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

historians, archivists, and digitization technicians, it was a perfect candidate to test linked open data metadata workflows and pilot a visualization of these agent relationships.

Metadata already performs a valuable role in describing what is known about collection materials, but too often it is not formed with the rigor necessary to be actionable in a disintermediated machine-readable environment. Even standards-based metadata can be subject to messy application of metadata guidelines, inconsistent maintenance of controlled vocabularies, or the partial or false matches of automated reconciliation services. In the SNJHP project, there are the additional challenges discussed previously including (1) metadata derived from different standards, with different user communities, (2) different systems, with different data structures, and (3) different levels of specificity. Despite these known issues, the team had agreed early in the development that people and groups would be key access points and that the digitization emphasis would be on creating sets of material such as digital objects, finding aids, and biographies, that centered around these agents. Despite a clear goal, there were missteps during the development of the project. One example was the inefficiency of having agent data stored in at least four locations without the ability to normalize it and easily transfer it between systems. This shortcoming reinforced the need to (1) recognize the overall data stream of the project and think beyond the smaller datasets or silos of information and (2) contribute "clean" or normalized data back into the upstream data sources in future projects. This was of particular importance when the metadata was transformed into Resource Description Framework (RDF), a semantic web standard that facilitates the interoperability of data on the web and the interlinking of data with assigned URIs.

The UNLV Linked Open Data Project had previously experimented with two types of visualization tools, Virtuoso PivotViewer and Relfinder. Both of these applications act upon the output of LOD transformation, RDF datasets, and provide access to the data via the semantic RDF query language, SPARQL. For the SNJHP, the UNLV LOD team conceptualized and developed a custom interface that optimized limited programming resources and focused on small scope goals. Rather than working to customize an existing product, the developer created an application to meet the stated goal of visualizing agent data. The functional requirements of the interface were to express three classes of the Europeana Data Model adopted for the project data transformation (1) Agent, (2) Information Resource, and (3) Provided Cultural Heritage Object (CHO). In this project, Agents refer to people and groups, the Information Resource refers to the original collection archival finding aid in portable document format (PDF), and the Provided CHO refers to the digital object. These three classes make up

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

the subjects and objects of the linked open data triples; two ontologies, Relationship and AgReLon, make up the predicates expressing the relationships among Agents. With this structure in place, focus turned to harmonizing the interface to match with the existing Drupal web content for the SNJHP, providing image files for the LOD interface, and evaluating designs including graph layouts, tile layouts, and various dynamic ways of refreshing the data as users navigate the linked data graph.

Future research will evaluate the LOD interface. Evaluation of LOD from the user perspective has not been explored deeply in the published literature, but as LOD projects mature, innovative user studies should be conducted to better document the impact of LOD-based information-seeking. For now, perhaps much of the value in researching LOD comes from how it can inform current practice and drive exploration into making metadata workflows more future-oriented. As the team looks forward to an online world with great potential for interconnected datasets, practical approaches to merging metadata and developing workflows that optimize the quality of data streams that originate in cultural heritage institutions will increase in importance.

Conclusions

The SNJHP made evident the need for cross-division metadata management and interoperability. The process of merging metadata for the SNJHP presented several challenges resulting from the differences of working with full legacy collections, prioritized curated collections, born-digital materials, newly donated collections, and scan-and-return collections. The audience for the SNJHP digital collection is primarily community users and donors, though secondary audiences included archival researchers, faculty, and K-12 educators. The unique nature of the project, with its focus on communities, people, and groups, led to interesting decision points about the value of metadata, as well as best practices for leveraging metadata to highlight and reveal connections between participants in the project and the larger community.

The following modifications, undertaken during the project and now being cascaded across workflows, have greatly benefited cross-division metadata creation and the potential for merging metadata in efficient workflows:

- implementation of ArchivesSpace collection management system;
- creation and discussion of Digitization Plans (which include an inventory of existing metadata and potential for reuse, decisions on granularity of

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

metadata approach, and selection of CV) prior to embarking on large scale projects;
- rigor in training and supervision of metadata creators to ensure compliance with the Application Profile and local procedures for LOD;
- adoption of principles guiding metadata approach: standards-based, transferable across systems, and user-focused.

In order to choose an appropriate metadata and access approach, all projects need to take into account the (1) overall goals, (2) research value of the collection in question, (3) time-frame available to complete the work, and (4) resources available to complete the project. Several competing interests are often present in grant-funded projects like the SNJHP. Generally speaking, all organizations can benefit from:
- conducting periodic inventories of the metadata characteristics present in their collection materials;
- reviewing how well this metadata travels between library systems where metadata is stored;
- working toward a culture of continuous improvement in metadata workflows to address both efficiency and quality benchmarks;
- and, keeping an eye to future methods of transforming this data into usable components for powerful machine-readable applications including linked open data.

There is no easy answer or silver bullet solution to merging and leveraging diverse metadata sets. By examining the history of digital collections and archival metadata development closely, as the authors did during the SNJHP project, much can be learned and institutional history can be understood and transcended.

References

Aikens, Barbara, Erin Kinhart, Robin Holladay Peak, and Judy Ng. 2011. "Digitizing Entire Collections at the Archives of American Art." *Archives of American Art*. https://www.aaa.si.edu/documentation.

Besser, Howard. 2004. "The Past, Present, and Future of Digital Libraries." In *A Companion to Digital Humanities*, edited by Susan Schreibman, Ray Siemens, John Unsworth. Oxford: Blackwell. http://www.digitalhumanities.org/companion.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

Bizer, Christian, Tom Heath, and Tim Berners-Lee. 2009. "Linked Data – The Story So Far." *International Journal on Semantic Web and Information Systems* 5 (3): 1–22. doi:10.4018/jswis.2009081901.

Dietz, Brian, and Jason Ronallo. 2011. "Automating a Digital Special Collections Workflow through Iterative Development." *Paper presented at the Association of College & Research Libraries Annual Conference*, Philadelphia, Pennsylvania, March 30-April 2. www.ala.org/acrl/sites/ala.org.acrl/files/content/conferences/confsandpreconfs/national/2011/papers/automating_digital_s.pdf.

Greene, Mark A. and Dennis Meissner. 2005. "More Product, Less Process: Revamping Traditional Archival Processing." American Archivist 68 (2): 208–263. doi:10.17723/aarc.68.2.c741823776k65863.

OpenLink Software. 2016. "Visualizing Your Data with PivotViewer." http://vos.openlinksw.com/owiki/wiki/VOS/VirtVisualizeWithPivotViewer.

UNLV University Libraries. 2015. *Southern Nevada Jewish Heritage Project*. http://digital.library.unlv.edu/jewishheritage.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

Ivey Glendon

14 Evolution and Revolution in Metadata Librarianship: Identifying New Requirements for Professional Qualifications amid Organizational Change

Introduction

The role of metadata librarian has been both ambiguously defined and consistently considered an evolving position since its introduction as a position in the library and information profession in the late 1990s. What is meant by the term "metadata librarian"? Is this professional just a cataloger with a new name? Or a programmer? Something in between? Or something entirely different? Job postings for metadata librarian positions vary greatly in the array of skills and qualifications sought for the position of metadata librarian in academic institutions, those with traditional cataloging skills, experience with digital initiatives, and scripting and coding expertise will all meet the qualifications for a metadata librarian position.

At the University of Virginia (U.Va.) Library, the position of metadata librarian has similarly been considered and reconsidered anew amid great organizational change in the Library. The Library hired its first metadata librarian in 2011 as the "non-MARC" (or MAchine-Readable Catalog) expert into what was then called the Cataloging and Metadata Services unit with the goal of accelerating access to digital collections. Since that time, the library has hired additional metadata staff, overseen departmental reorganization, and undergone Library-wide reorganization in response to university-level changes in leadership, higher education, and new strategic directions.

The purpose of this chapter is to analyze metadata librarianship broadly in the profession as well as specifically within the context of major organizational change at the University of Virginia. Metadata work has been happening at the U.Va. Library for more than a decade. This chapter will demonstrate how the library's newest conception of organizational design around metadata work has enabled significant accomplishments, revealed both challenges and opportunities with regard to staffing, and represents demonstrable advancement in the Library's ability to meet new and changing university directives.

https://doi.org/10.1515/9783110337419-015

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

Defining metadata librarianship and its role in an academic library

Unlike other, more widely understood areas of librarianship where reference librarians interact with users at a service point and where subject librarians provide robust and specialized research expertise, the role of the metadata librarian resists consistent and agreed-upon characterization. Literature in the last decade has deemed the position of metadata librarian as the digitally minded sibling to the cataloging librarian responsible for description of print collections. In 2007, John Chapman (2007) noted in his analysis of the role that "many institutions have placed the position of metadata librarian within a traditional cataloging or technical services department" (279) and that in doing so they aim to use the position as a "fulcrum" to balance the traditional cataloging skills with new digitization initiatives. Likewise, Han and Hswe (2010) observed in their review of metadata librarian job announcements spanning a nine-year period that the position is (1) usually located in a technical services unit, (2) often has experience with traditional, item-level cataloging, and (3) that the main distinction between cataloging librarianship and metadata librarianship is the expectation that the latter position "have knowledge of emerging technologies" including "markup languages such as XML, protocols such as the OAI-PMH, and approaches to conceptual modeling such as RDF" (Han and Hswe 2010, 137).

At the same time, Eden (2010) accepts that metadata librarians may have a natural home in technical services units but argues that the work is not nearly as overlapping or related as Chapman (2007) and Han and Hswe (2010) would suggest. Additionally, Eden (2010) noted that "the status quo of technical services is no longer viable or cost-effective; all of us must...restructure our organizations to collaborate and consult with users regarding their information and research needs" (99). According to Ma (2007), the Association of Research Libraries (ARL) supports Eden's (2010) thesis noting that at most ARL institutions "the metadata librarian plays a leadership role in metadata activities while... working with systems personnel, subject specialists, project partners, and even end-users on metadata-related issues; documenting metadata policies, procedures and guidelines; and training staff" (Ma 2007, 15). Additionally, according to Ma (2007) about 85 percent of respondents (65 institutions) to the ARL survey also noted that their organizations had changed to provide metadata services, with more than half of respondents indicating they had redefined positions and had distributed metadata activities across the institution.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

At the time of these analyses, metadata librarians were developing their own communities, sometimes within technical services departments but also alongside cultural heritage professionals and library programmers. Within the American Library Association (ALA), the Metadata Interest Group became a formal entity in the Association for Library Collections and Technical Services Division in 2009 (American Library Association 2016). Other online communities, such as Code4Lib, provide opportunities for metadata librarians to engage with self-described "hackers, designers, architects, curators, catalogers, artists and instigators from around the world" (Code4Lib 2016, under "Summary"). The professional communities for metadata librarians are incredibly diverse in terms of skills and interests, mirroring any set of contemporary job announcements for the position.

It is clear from literature on metadata librarianship and in documentation from the metadata librarian communities that the past decade has been a varied experience in the profession for those involved with this work. It is also obvious that the lack of consensus on what the role of the metadata librarian is and where it should function in a library reflects the highly decentralized nature of metadata librarianship. After all, metadata services power all major functions of discovery, access, and preservation of library-managed content. The history of metadata work and metadata librarianship at the U.Va. Library has identical origins and is the basis for understanding how the library has arrived at its contemporary understanding of metadata positions and organizational structure.

Evolving notions of metadata librarianship in a changing organizational context

The U.Va. Library has a deep and rich history with regard to its collections, services, and digital initiatives. Fifteen libraries comprise the library system and shares close ties with its "sister" libraries at the university's professional schools (School of Medicine, School of Law, and Darden School of Business). In 2015, the U.Va. Library counted among its collections 5.3 million print volumes, seventeen million archives and manuscripts items, and 22.1 million digital objects (University of Virginia Library 2016). The Library began collection digitization in the 1990s, and has been a leader in library technology (including co-developing the repository software Hydra), as well as a partner for institutional research units such as the Sciences, Humanities, and Arts Network of Techno-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

logical Initiatives (SHANTI) and the Institute for Advanced Technology in the Humanities (IATH).

Committee-based metadata work began at the library in 2003 in the form of a Metadata Steering Group (MSG), which was responsible for the approval, and maintenance of metadata standards for the Library's digital initiatives. Its membership reflected a decentralized, location-based, and discipline-centric approach to the task, with members representing (1) cataloging services, (2) the fine arts library, (3) the science and engineering library, (4) the music library, and (4) digital research instruction and development teams. Available minutes for the MSG reflect a team active from fall 2003 to summer 2007.

Personnel changes in 2008 marked a shift in the library's approach to metadata. A new Head of Cataloging Services joined the organization and the person's prior role as a metadata librarian at another institution informed new strategic directions for cataloging services at the Library. Due to an uncertain economic climate and a university-wide slowdown on hiring, few positions were advertised at the Library following the recession in 2008. The position of metadata librarian, the Library's first such role, was advertised and filled in 2011 when hiring again became a possibility.

The hiring of the new cataloging manager and the first metadata librarian represents the beginning of a new era of metadata services and librarianship; a period marked by (1) continuous departmental and organization-wide change, (2) refinement in required knowledge, skills, and abilities for metadata positions, and (3) new articulation of the skills necessary to achieve success in this work. Between 2010 and 2015, the cataloging services unit reorganized multiple times, resulting in an evolution that reflects (1) an increasing centralization of metadata services, (2) the need to expose staff to the concept of change, and (3) a library-wide shift focusing on services by function rather than by discipline or location.

Between 2010 and 2015, the department formerly known as "Cataloging" adjusted its mission twice to expand its scope, responsibility, and services available to library staff and researchers in the university. The unit made an intentional pivot to be more externally facing to library colleagues and to researchers at the university, many of whom were interacting with metadata for the first time as a result of new requirements from granting agencies such as the National Science Foundation (NSF) and the National Institutes of Health (NIH). During this time, metadata staff greatly increased the frequency of their engagement with faculty, participated in more research initiatives, and conducted more training.

The Library's full reorganization in 2015 validated the concepts of the previous departmental reorganizations and brought additional possibility to meta-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

data services. There was a new "Acquisitions and Discover Department" created which housed three sub-units: (1) Resource Acquisition and Description, (2) Metadata Creation and Organization, and (3) Metadata Analysis and Design. Within this organizational arrangement the Acquisitions and Discovery Department has responsibility for and oversight of all metadata for all library-managed collections. This represents a turning point in the Library because it represents a move away from our fragmented and decentralized past and toward a more holistically arranged organizational structure. The library's reorganization of the Acquisitions and Discover Department in 2015 provided the opportunity to strategically assess the essential roles within the Departments three sub-units.

Another interesting approach to redefining the metadata librarianship at the U.Va. Library was the creation of a "Metadata Analysis and Design" division where (1) the Archivist, (2) the Metadata Librarian for Research and Scholarship, and (3) the Metadata Librarian positions directly supported an integrated and holistic approach to metadata management and consultation throughout the Library. These positions require an understanding of multiple metadata formats in order to meet researchers' needs.

The resulting effect of the U.Va. Library reorganization have allowed leaders and managers in the Library to develop new conceptions of the positions and the soft-skills necessary for success within metadata librarianship. Because of the scope of responsibility, the highly collaborative work, and equal focus on internal and external clients, these positions require positive interpersonal skills, a commitment to relationship building, and a curious and flexible spirit; traits that have been important in the past but which are more essential than ever.

The intersection of metadata librarianship and university-level strategic directions

It is important that those involved with metadata librarianship and metadata work in the library also understand how metadata services impact and support strategic directions at the University-level. In 2013, the university announced a new strategic mission entitled "The Cornerstone Plan" which is a "common framework" intended to guide leaders throughout the institution (University of Virginia 2015). Five strategic "pillars" comprise the Cornerstone Plan:

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

- enrich and strengthen U.Va.'s distinctive residential culture
- advance knowledge and serve the public through research, scholarship, arts and innovation
- provide educational experiences that deliver new levels of student engagement
- assemble and support a distinguishing faculty
- steward resources to promote excellence and affordable access

The Library has an opportunity to positively influence each of these five strategic pillars and metadata librarianship will continue to take an active role in advancing the library's capacity to productively contribute to them. For example, because of our metadata staff's commitment to providing expertise throughout the university community, we have been able to develop relationships with different organizations and act as metadata consults. The U.Va. Library also has an ongoing and productive relationships with faculty across the institution, particularly in the School of Architecture and in the College of Arts and Sciences. Finally, our renewed emphasis on rare and unique materials as a result of the integration of Special Collections and non-special collections metadata expertise positions us well to positively impact all Five Pillars of the new strategic direction.

Metadata librarianship is active in all functions of the library such as (1) the search, discovery, and access for all materials in all formats, (2) user experience both in-person and online, (3) metadata support for individual research, and (4) the development of preservation metadata strategies. With the creation of the university's new strategic directions, metadata librarians at U.Va. are now moving away from the notion of a "back-of-the-house" technical services operation and are embracing the concept of providing agile and mission-driven metadata services for students and faculty within the university.

New understandings of metadata librarianship and smart organizational design

As noted, the Acquisitions and Discovery Division is comprised of three metadata sub-units: (1) Resource Acquisition and Description, (2) Metadata Creation and Organization, and (3) Metadata Analysis and Design. This centralized configuration of metadata services has generated several points of success in (1) connecting patrons to resources, (2) integrating with Special Collections meta-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

data, and (3) collaborating with colleagues throughout the Library in order to meet the Five Strategic Pillars outlined in the mission of the university. Additionally, the centralized configuration of the Library's metadata services have reduced the number of cataloging errors. Finally, in addition to the benefits noted above, the managers of the Acquisition and the Discovery Division placed the function of data integrity with Metadata Analysis and Design, thus separating metadata creation and correction thereby creating a neutral space for metadata corrections to occur; this means that metadata creators no longer correct their own work.

Three positive results are also reflected in the centralization of "Special Collections" and "non-special collections" under the Metadata and Analysis subunit: (1) implementation of ArchivesSpace, (2) development of a new software called "Transmog," and (3) development of a plan to programmatically convert legacy metadata into a modern schema. Adoption of ArchivesSpace in 2015, enabled the Library to move from pre-web paper finding aids to the digital world. This has been an enormous challenge as there are 15,000 archival collections. In working through this challenge, metadata staff have collaborated with a software developer to build "Transmog," which is a web-based tool that uploads and "reads" finding aids in Microsoft Word format to detect patterns such as scope and content notes, and converts them to valid Encoded Archival Description (EAD) eXtensible Markup Language (XML) format. Projects such as this would be impossible were it not for the U.Va. Library reorganization.

The new centralized organizational structure has also improved metadata services throughout the Library. Unified under the Acquisitions and Discovery Division, each sub-unit is better able to respond to collections issues and present a more holistic approach to workflow plans when assessing collection acquisition. In one such case, the Division was able to work together with the Library's development office to assess a donor's major collection in support of new curricular initiatives and provide timely cost estimates for acquisition and processing of materials. As a result, the Division provides a structural organization that offers a smarter and more responsive service to our colleagues in the collection management areas.

Beyond collections, centralization of metadata Division has had a positive impact on a variety of metadata services. With metadata staff workloads balanced in three separate units, metadata staff, particularly those in Metadata Analysis and Design, have been able to forge closer relationships with staff and digitization services. Additionally, metadata librarians are better able to consider user experience issues such as how metadata impacts faceted display and provide metadata management expertise for print collections.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

The centralization of metadata services in the Acquisitions and Discovery Division has greatly increased the capacity of metadata librarians at the U.Va. Library to provide high impact services in regard to research, faculty, and increase the quality student experience. Prior to the reorganization, metadata expertise was highly fragmented which led to inefficient communication, slower response times to students and researchers, and fractured consultation that lacked a broader view of relevant issues. By centralizing metadata staff around the functions of metadata librarianship, the U.Va. Library has created an environment where the organization can operate as a thriving, innovative laboratory for new ideas and strategic approaches for metadata management.

Conclusion

The U.Va. Library's evolution in metadata librarianship has afforded the opportunity to identify challenges, opportunities, and new potential in moving forward as a more agile organization. In addition to the usual challenges associated with organizational change, institutional leaders have realized that there are other cultural issues that influence the organization of the Library. Some staff involved with metadata work demonstrate intense care for local practices, possess a strong affinity for meticulous item-level description, and are content with traditional workflows and methods for metadata creation and management, which are characteristics that well served the library of the past. These traits may not be unique to this institution and may instead reflect the rule and process-based nature of metadata work; however, put together they add up to a slow pace of change when the environment is moving much more quickly in academic libraries. The challenge now is to cultivate norms around resilience in change (as opposed to simply coping with change) and support increased interaction with professional communities.

At the same time, the library's reorganization in 2015 has injected new life and possibility into metadata librarianship at the library. Also, the new staff configuration supports building on the library's record of accomplishment while simultaneously supplanting cultural threats with new and positive norms. In addition to the optimism engendered by new organizational design, making strategic hires into metadata services will be critical to the library's success in the future. This is especially important when creating positions to advance the library's mission of increased openness to collections and their metadata. To achieve this, the organization will need to hire individuals who have the adapt-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

ability, teamwork, project management, and communication skills to peer into tomorrow and advance the library into the web 3.0 environment.

This is a truly exciting time in metadata librarianship and higher education. The library is well-positioned to make collections accessible more quickly and to provide metadata expertise to the university's researchers faster and more responsively than before by fully reimagining metadata services and deploying metadata expertise strategically across the library and university. With the organizational structure in its smartest design yet, metadata librarians at the U. Va. Library are poised to engage with colleagues locally, regionally, and nationally to participate fully in conceiving and implementing standards and processes to ensure access to and sharing of library collections and university research as widely as possible in a rapidly shifting research ecosystem.

References

Chapman, John W. 2007. "The Roles of the Metadata Librarian in a Research Library." *Library Resources and Technical Services* 51 (4): 279–285. doi:10.5860/lrts.51n4.279.

Code4Lib. 2016. "About." http://code4lib.org/about

Eden, Bradford Lee. 2010. "The New User Environment: The End of Technical Services?" *Information Technology and Libraries* 29 (2): 93–100. doi:10.6017/ital.v29i2.3148.

Han, Myung-Ja, and Patricia Hswe. 2010. "The Evolving Role of the Metadata Librarian: Competencies Found in Job Descriptions." *Library Resources and Technical Services* 54 (3): 129–141.

Ma, Jin. 2007. *SPEC Kit 298: Metadata*. Washington, DC: Association of Research Libraries.

University of Virginia. 2015. "Current Strategic Plan." http://planning.virginia.edu/about/current-strategic-plan.

University of Virginia Library. 2016. "Size of Collections." http://assessment.library.virginia.edu/files/2016/01/2015CollectionsRev.pdf.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

Jocelyn Olson Triplett

15 Increasing Interoperability through the Transformation and the Consolidation of Image Collections' Metadata

Introduction

The ultimate goal for many of us who manage digital collections is to create collections that are interoperable, described with rich, standard, and shareable metadata, and paired with a technological infrastructure that enables easy discovery and use.

The University of Virginia (U.Va.) Library has been making digital image collections available to the public for nearly two decades, however the consolidation of our digitized photography collections into one repository architecture and method of description is only now nearing completion. In this chapter the author discusses challenges encountered while standardizing and migrating several varieties of metadata, the technological architecture implemented by the U.Va. Library for dissemination of the collections, and how decisions made during the migration process affected the discovery and display of the collections in Virgo, U.Va.'s online public access catalog.

Literature review

The interoperability of digital collections is dependent upon the quality of both its resources' description and metadata. In order to create usable digital collections, the most important task for librarians is to ensure that the collections' metadata quality is "good," as "if metadata quality is poor so is the discovery of digital library information objects" (Tani, Candela, and Castelli 2013, 1195).

What "good" and "poor" mean, however, varies according to the goals of the institution for how the collections are to be discovered at the time of metadata creation. Inevitably what was once "good" metadata becomes "poor" as an institution's digital library technologies or goals for its digital collections change. Consequently, it is necessary for institutions to periodically evaluate whether their metadata meets the criteria set forth by the National Information Standards Organization (NISO) *A Framework of Guidance for Building Good Digi-*

https://doi.org/10.1515/9783110337419-016

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

tal Collections as in its first metadata principle: "Good metadata conforms to community standards in a way that is appropriate to the materials in the collection, users of the collection, and current and potential future uses of the collection" (NISO 2007, 61).

As aggregation across collections and institutions has become more and more common over the previous decade, meeting the criterion of NISO's second principle, "Good metadata supports interoperability" (NISO 2007, 76), has become essential for many institutions desiring "good" metadata.

Interoperability is the "compatibility of two or more systems such that they can exchange information and data and can use the exchanged information and data without any special manipulation" (Taylor and Joudrey 2004, 369). It can refer to the compatibility of collections within a collection, repository, or institution, but commonly for digital libraries, "the goal of interoperability is to help users find and access information objects that are distributed across domains and institutions" (NISO 2007, 77). The primary benefit for institutions is that "by allowing their metadata to appear in places outside of the original collection, institutions increase the number of access points to the items in their collection and expose their collection to a broader audience" (Shreeves, Riley, and Milewicz 2006, 2). For collections that are shared, in order for them to be accessible and meaningful, the responsibilities fall on both the organizations that create the metadata records and make them available, and the service providers who aggregate the collections (NISO 2007, 76). For the data providers, simply being able to share their collections is not enough; "While sharing metadata is an essential first step towards creating useful aggregations, the quality of the resulting services is limited when the metadata used is not interoperable" (Shreeves, Riley, and Milewicz 2006, 3). For those institutions that discover their metadata is no longer (or never was) sufficient, their staff is soon faced with the task of enhancing, migrating, or mapping their metadata to ensure that their digital collections' metadata quality is both "good" and the collections are interoperable.

Park and Tosaka (2010) state that "metadata interoperability is among the top challenges faced by digital libraries" for which the quality of the metadata is "critical" (104). For institutions that want to increase the interoperability of their collections, determining the current and desired quality of their metadata collections individually is the first order of business, followed by a plan for accomplishing the necessary changes to make up any differences. Yasser (2011) identified five different categories of problems in digital library metadata that can affect the quality of the records. The first, "incorrect values" (60) is "where elements are applied according to the prescribed semantics of the scheme, but the associated values do not describe the resource accurately" (60). "Incorrect

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

elements" (60) "are not applied according to the scheme." "Missing information" (60) happens when elements are not present. "Information loss" (60) occurs when attributes are mapped to data elements whose semantics are less granular. Finally, "inconsistent value representation" (60) occurs when the values of the element are not recorded in consistent forms (Yasser 2011, 60). At a base level, "while the quality and richness of metadata supplied may vary, as long as it has been applied in a somewhat consistent manner, it can be repurposed and enhanced in order to create a more complete metadata record" (Corrado and Jaffe 2014, 44).

There are a number of methods one can use to increase the quality of metadata descriptions depending upon the problems identified. Tani, Candela, and Castelli (2013) believe that these different methods can be grouped into four categories of approaches, noting however that none of the methods individually "can be considered as the ultimate and optimal solution to all quality issues, especially in complex and heterogeneous contexts" (1202). Approaches described by Tani et al. (2013) that might be taken are those that achieve a common understanding of metadata (applying guidelines, standards, and application profiles), those that highlight the problems affecting the metadata (evaluation methods), those that aim at supporting the generation of metadata (semi-automatic metadata creation), and those that aim at repairing and homogenizing metadata (normalization, enhancement, and augmentation).

When the quality of a collections' metadata has been addressed, the next step is determining the level of interoperability needed for the collection, which may spur additional metadata changes. Existing issues that had gone unnoticed in earlier uses of the collections are "further amplified when dealing with 'big data' scenarios where data and metadata (1) come from multiple and heterogeneous sources, (2) are collected with different approaches, and (3) are expected to be used in contexts different from their initial ones" (Tani, Candela, and Castelli 2013, 1203). Depending on the source of the problems, efforts to improve interoperability may need to happen at the schema, record, or repository level, or a combination of the three. Chan and Zeng (2006) describe the results that can be observed after improvements at each of the three levels:

- Schema level: efforts are focused on the elements of the schemas, being independent of any applications. The results usually appear as derived element sets or encoded schemas, crosswalks, application profiles, and element registries.
- Record level: efforts are intended to integrate the metadata records through the mapping of the elements according to the semantic meanings of these elements. Common results include converted records and new records resulting from combining values of existing records.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

– Repository level: with harvested or integrated records from varying sources, efforts at this level focus on mapping value strings associated with particular elements (e.g., terms associated with *subject* or *format* elements). The results enable cross-collection searching. (Chan and Zeng 2006, first page)

For collections that had been described in silos, the context was often a large part of understanding the resources. The problems that occur when the context is removed are common: "Digital collections with a topical focus are notorious for creating non-interoperable metadata when they assume that users know the main topic of the collection. When this metadata is shared in larger aggregations, descriptions that made sense in the context of the original collection can be mystifying" (NISO 2007, 76). This metadata may need to be addressed at the record level, with additional metadata elements added or values changed to make the context clear.

For legacy collections, the method of resource description was often chosen for reasons that did not include interoperability. In their survey of cataloging and metadata professionals, Park and Tosaka (2010) discovered that "the leading criteria in selecting metadata and controlled vocabulary schemata are derived from collection-specific considerations of the type of resources, the nature of the collections, and the needs of primary users and communities" (114). For these collections, which may have rich and consistent description, but perhaps no formal schema, or a schema that does not support the level of interoperability desired, increasing interoperability will require migrating or mapping to a community-supported schema.

Increasing the interoperability by migrating to a new schema often has tradeoffs. When moving to a new schema, catalogers should be aware that "the major challenge in converting records prepared according to a particular metadata scheme into records based on another schema is how to minimize loss or distortion of data" (Zeng and Chan 2006, first page). As Attig, Copeland, and Pelikan (2004) discovered when working with their own data, "the authors found this warning to be amply warranted. There were semantic differences between the definitions of superficially similar data elements" (255). And similarly, de Groat (2009) in her discussion on metadata remediation for aggregators notes "records that have been mapped from legacy data in other formats will seldom be optimized in their new home, and the creators of these records may not have the resources to augment these records in any more than the simplest ways" (2).

The last level for examining interoperability is at the repository level. At this level, "an issue often encountered is that the participants may have used diverse schemas and description methods to create their metadata records" (Chan

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

and Zeng 2006, first page). A system must be developed that can manage multiple schemas and mappings in order to enable cross-collection searching. Since it is essentially the same process as mapping between schemas, the same drawbacks can occur at the repository level mappings of loss or distortion of data. For example, users may need to search for a resource in a "creator" field that was mapped from metadata describing the value of the field as the more specific "author" "painter" or "composer."

Digital Image Collections in the University of Virginia Library

Many of the types of problems described by Yasser (2011) occurred in the metadata for the University of Virginia's digital collections. Digital image collections have been part of the University of Virginia Library's collections building efforts for the last twenty years. The U.Va. Library was an early leader in bringing technology into the library, establishing the Electronic Text Center in 1992 to build extensive collections of Internet-accessible digital texts. Four years later, in 1996, the library embarked on its first large, donor-funded image digitization project, when the Holsinger Studio Digital Portfolio was created, making an online database of 10,685 photographs available to the public. Depicting local Charlottesville, Virginia scenes as well as studio portraits, the images, which had been captured on plate glass negatives, were rapidly deteriorating. Metadata for the digital portfolio was derived from the photographer's original studio ledgers. Staff had first transcribed the metadata into a dBase database, before migrating it into Microsoft Access and then finally into FileMaker Pro, which provided the metadata for the online database. This online database presented the viewer with six types of information for each image: (1) name (the customer who purchased the photograph), (2) date, (3) photographer's negative number, (4) free text comments, (5) size of the negative, and (6) one of five categories identified by Library staff. Subsequent to the success of the Holsinger collection, the Library continued to create more collections according to this model into the mid-2000s.

In 1999, the U.Va. Library became involved in digital repository development and chose Fedora as its content management framework. Originally developed at Cornell University as a research project, the U.Va. Library became a collaborator in developing the open-source Flexible and Extensible Digital Object Repository Architecture.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

With the implementation of the Library's first production repository in the early 2000s, the first efforts to begin consolidating the Library's image collections into Fedora began. As part of this effort, the Library created several local eXtensible Markup Language (XML) standards in a move towards standardizing digital object metadata including (1) the General Descriptive Modeling Scheme (GDMS), (2) U.Va. DescMeta, and (2) U.Va. AdminMeta. GDMS was envisioned as a "formal information structure that [could] be used to construct descriptive models ... to create contexts for collections of digital resources" (University of Virginia Library Digital Initiatives 2011, first page). It was the Library's initial attempt at standardizing metadata for inclusion in the repository, and several collections' metadata were migrated into GDMS and their derivative images made available through Virgo, the U.Va. Library's online public access catalog.

GDMS, U.Va. DescMeta, and U.Va. AdminMeta were abandoned near the end of the decade in favor of adopting a national standard. The Library committed to adopting the Metadata Object Description Schema (MODS) standard for its image collections shortly after the release of the Digital Library Federation/Aquifer *Implementation Guidelines for Shareable MODS Records* in 2006.

Also under development in the mid-2000s by the U.Va. Library was a Ruby on Rails application called Tracksys, which managed the Library's special collections digitization workflows. In 2009, work began to migrate all digital collections through Tracksys, which accomplished several objectives: all preservation Tagged Image File Formats (TIFFs) were moved to one storage location and duplicated to redundant tape storage. Technical and preservation metadata were created during the ingest process. Finally, repository objects, access derivatives, Solr index records, and Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) Dublin Core (DC) were created for each image entering the digital library. Through management in Tracksys, it became possible for U.Va. Library staff to update descriptive metadata and download image files using a graphical user interface.

The reward for completing these collection migrations is the consolidation of all the metadata for image collections into a standard metadata format and repository, and the archival files into a single environment. This makes it easier to maintain the metadata going forward, either for updating the MODS standard or moving to a different one in the future. Metadata edits for all the collections can now be done from a single interface, and the migration process created checksums and other preservation technical metadata necessary for monitoring the TIFF files.

This chapter will review three of those collections migrated into Fedora through Tracksys, namely (1) the Jackson Davis Collection of African American Educational Photographs, (2) the University of Virginia Visual History Collec-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

tion, and (3) the Frances Benjamin Johnston Photographic Collection. Although these digital collections were developed within a relatively small span of time, the late 1990s through mid-2000s, it was a time of rapid development of metadata standards and digital repository technology in the Library. Consequently, each required different types of metadata normalization or enhancement for migration into what are now our digital library collections accessible through Virgo.

Migration process

Each of the collections, regardless of the original method of description, followed a similar process for migration into Tracksys and then Fedora. The following steps were taken:

- evaluation of the collection's metadata
- normalization or enhancement of metadata values, if needed
- creation of a crosswalk for the collection's metadata to MODS
- creation of an Extensible Stylesheet Language Transformation (XSLT) of the metadata using the <oXygen/> XML text editor
- scripting of a batch upload of the newly created MODS records and corresponding image TIFFs into Tracksys
- automatic creation of Solr index records, a Dublin Core metadata record, Metadata for Images in XML Standard (MIX), and JP2K derivative images by Tracksys, which then pushes these along with the MODS records into Fedora

Visual History

The University of Virginia Visual History Collection, created in 2003, consists of around 16,000 images depicting the people, architecture, and events associated with the University. The collection grew from the digitization of the Special Collections Library's "prints file" which contains visual materials that had been pulled from their original manuscript collections and aggregated into one location. Eventually, several areas contributed images and metadata across the University, including the Health Sciences and Law Libraries, Alumni Association, and Media Relations Office.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

The Visual History Collection was cataloged in a customized FileMaker database, with categories based on the arrangement of the physical materials, grouped together under headings like "portraits," "athletics," or "faculty/student." Other fields available to the image catalogers included "title," "date," "accession number," "artist/photographer," "format," "comments," and "restrictions," enabling fairly rich description of the materials. While guidance was provided for the type of information to be included in each field, there were no required fields (aside from an image identifier) or required formatting, and the distributed nature of the record creation resulted in some inconsistent data, and consequently some headaches during the migration process.

The biggest problems encountered in the metadata were the inconsistently formatted dates and creators, and lack of information about the copyright status of the items. Catalogers had been advised to write the date of the item based on information recorded on the item or its folder. This resulted in a wide variety of formats, for example (1) Orig.1856; 1934, (2) early 50s, (3) Civil War, (4) Apr-07, (5) Post – 1895 [?], (6) ca 1917, (7) ca. 1835–1837, (8) circa 1948; c.a.1900, (9) c.1860, (10) n.d., (11) nd, (12) April, (13) Unrecorded, and (14) unknown. With an end goal of having the dates in the World Wide Web Consortium Date Time Format (W3CDTF) "YYYY-MM-DD" as recommended by the *MODS Implementation Guidelines*, all the dates were first standardized to the most common format within the database, MM-DD-YYYY or date range of MM-DD-YYYY – MM-DD-YYYY. These were then transformed to W3CDTF format programmatically. The uncertainty conveyed through brackets, question marks, "circa" or other means became "inferred" or "approximate" values on the dateCreated qualifier attribute.

While normalizing the dates had been tedious, the most daunting part of the project was enhancing the records with standard, authorized names for the image creator field. Fortunately, just as the migration was getting underway, a volunteer contacted the Library looking for a project that would be appropriate for working on remotely. The image identifiers and creator fields were exported to a Google spreadsheet, and the volunteer was able to work through the list, searching the Library of Congress Name Authority File (LCNAF) and updating the spreadsheet as she found matches.

To create the final MODS records, the FileMaker metadata was exported as an XML document, and then transformed into individual MODS records using an Extensible Stylesheet Language (XSL). A second XSL transformation was run that inserted the creator name that had been determined by the volunteer into each new MODS record.

Although there had been a field in the FileMaker database for indicating images that were believed to have copyright restrictions associated with them, the

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

field was not used reliably, and there was no way to do a programmatic assessment of all the images to determine which were free from copyright. A Library archivist did a cursory evaluation and indicated the identifiers for those images he believed should be withheld from the public catalog until an in-depth evaluation could be done, and those images and metadata were ingested by Tracksys but were not put into Fedora.

The migration of the collection was largely a success; however, one problem arose when the entire collection was merged into the catalog that had not been foreseen. The original FileMaker records had a field for "format" that had been assigned forty different values. These had been mapped to the MODS "genre" element, and were displayed under "format" in the Virgo records. That facet had previously been limited to the more general terms "Book," "Map," "Periodical," etc., but after the migration now included "mezzotints" and "stipple engravings (prints)." The decision was made to consolidate the diverse Visual History formats into the Solr value for "visual materials," streamlining the catalog display, but disabled faceting by the specific formats that users had been accustomed to through the collection's previous online interface.

Jackson Davis Collection

The Jackson Davis Collection of African-American photographs contains nearly 6,000 photographs of African American schools, teachers, and students throughout the southeastern United States, taken between 1917 and 1947. The creation of the digital image portfolio was funded by a National Leadership Grant from the Institute of Museum and Library Services, and began in 1999. Images were digitized from nitrate film negatives, glass lantern slides, or occasionally photographic prints, if the original negative no longer existed.

The digitized images were cataloged in a FileMaker database and were described using data from Jackson Davis' own card catalog system. This information included the negative number, state and county in which the picture was taken, name of the school, and occasionally the date. As cataloging workflows evolved, the metadata went through two migrations before it was migrated to MODS in 2013. In 2004, the collection metadata was migrated into the Image Resource Information System (IRIS), a FileMaker cataloging utility based on Visual Resources Association Core Categories (VRA Core) and developed at Brown University. During its time in IRIS, the records were enhanced with subject headings and location information. At the time, the Library envisioned a workflow for digital image collection building that involved cataloging in IRIS, and

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

then an export of the IRIS metadata into GDMS for ingest into Fedora. After the metadata enhancements for the Jackson Davis collection were complete, this was what happened in 2007.

The 2013 migration of the Jackson Davis GDMS to MODS was the easiest of the three collections. The GDMS mapped easily, and the date format changes were accomplished programmatically. The collection had only one creator, and data cleanup had already been accomplished while in IRIS.

Frances Benjamin Johnston Photographic Collection

The final collection, the Frances Benjamin Johnston Photographic Collection, is comprised of nearly 1,000 images of colonial architecture in Virginia. Johnston, a Virginia native and one of the earliest female photographers and photojournalists, was active from the late nineteenth to the mid-twentieth centuries. Johnston captured the 948 photographs held by the University of Virginia Library between the years of 1933 to 1935 during a survey of Virginia architecture funded by a grant from the Carnegie Corporation (Gushee 2008, 10–11).

The original photographs, held by the Special Collections Library and depicting lesser-known structures of the Colonial Era, such as barns, log cabins, inns, and outbuildings, were photographed in the 1980s and duplicated in the form of a slide. These slides were then added to the Fine Arts Library's Image Collection, and were cataloged using MARC in the Library's online catalog. A project to make the images available online, supported by the Arts Council, digitized the original photographs between 2003 and 2006. The metadata describing the slide copies was exported from the Library's catalog to IRIS, and modified and enhanced with subject headings and location information. That metadata was then exported to XML and scripted using Apple Scripts and Purl into GDMS for inclusion in the digital repository. The collection's images were presented hierarchically in Virgo, maintaining the work/image relationship that is part of VRA Core. Discovery was at the work level, but because of technical limitations, much of the metadata at the image level was not displayed.

For migration into MODS, the decision was made to work with the metadata exported from IRIS instead of the GDMS metadata. Some metadata fields had been lost in the conversion to GDMS, and the exported IRIS metadata was easily viewable in a spreadsheet, making assessment of the contents of metadata fields across the collection easier, unlike the individual GDMS image records.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

The collection had many free text keywords in addition to properly formatted subject headings, which required significant cleanup. The keyword cleanup was performed using OpenRefine software, and these were then added to the MODS records as unauthorized subjects. A drawback to the migration was that the work/image relationship was lost as the database was flattened. We found, like Attig et al. (2004, 256) that our current digital library architecture does not handle the two-level VRA structure well, and does not allow us to both display a hierarchical representation of multiple images underneath one work record while still displaying detailed metadata for each image.

Conclusion

The digital image collection migrations that have been in progress in the U.Va. Library since 2009 have resulted in greater interoperability for our collections in several ways. By performing data cleanup as the first step of the migrations, the quality of the metadata was increased and the ability to discover relevant records within the collections was enhanced. For maintenance activities, the collections' records no longer need to be managed individually; it became possible to script updates and corrections across the collections as a whole. With only one metadata standard, the task of indexing the collections for discovery with the Library's other collections was also greatly simplified. By choosing a nationally supported metadata schema, and consolidating the collections' metadata into one repository collection, we enabled the aggregation of our records by the Digital Public Library of America (DPLA), making them interoperable with collections across other institutions.

Because of the different creation histories each collection migration was a learning process. As the collection migrations near completion, the collections would be well served by a holistic review of the digital image collections, as well as a review collection-by-collection in order to evaluate whether some of the unique metadata can be highlighted within the collection without hurting interoperability. During the migration process, the work emphasized applying the implementation guidelines towards making quality MODS records, but there was less effort put into making sure that the MODS were consistent between collections. Corrado and Jaffe (2014) described the work they did mapping collection metadata keywords to authorized subjects: "While this process initially seems like a lot of work and some may question the value of it, given the volume of photographs the Libraries are processing, it can be viewed as an investment" (40). With the volume of MODS records in the U.Va. repository greatly

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

increased, it may be worth a similar investment to map unauthorized subjects and other keywords within the collections to Library of Congress Subject Headings (LCSH) in order to make the collections more interoperable with the other resources available through the catalog, and to create scripts or other tools based on the migrations already completed that can be adapted towards future migrations.

A recent project added Vanity Fair Caricatures to the Library's digital collections, and the Virgo developers were able to make customized faceting within the collection so that users could facet by categories that were unique to the collection. As a recent enhancement to the digital library, it would be worthwhile to review the other image collections to see if the feature could be added to take advantage of metadata that is not common across the other collections.

As we approach the end of the migration work needed to add the Library's legacy collections to our current repository architecture, it is important to look forward to new possibilities for using the collections' metadata, and continually reevaluate our collections for quality and interoperability in the changing digital library landscape.

References

Attig, John, Ann Copeland, and Michael Pelikan. 2004. "Context and Meaning: The Challenges of Metadata for a Digital Image Library with the University." *College & Research Libraries* 65 (3): 251–261. doi:10.5860/crl.65.3.251.

Chan, Lois Mai, and Marcia Lei Zeng. 2006. "Metadata Interoperability and Standardization – A Study of Methodology Part I." *D-Lib Magazine*, 12 (6). doi:10.1045/june2006-chan.

Corrado, Edward M., and Rachel Jaffe. 2014. "Transforming and Enhancing Metadata for Enduser Discovery: A Case Study." *Journal of Library and Information Science* 5 (2): 34–47. doi:10.4403/jlis.it-10069.

de Groat, Greta. 2009. "Future Directions in Metadata Remediation for Metadata Aggregators." Digital Library Federation. https://old.diglib.org/aquifer/dlf110.pdf.

Digital Library Federation. 2009. *Digital Library Federation/Aquifer Implementation Guidelines for Shareable MODS Records. Version 1.1.* https://www.semanticscholar.org/paper/Digital-Library-Federation-Aquifer-Implementation-Shreeves-Shawver/b7c27d3728470a4774a1bb597a79f15df60c106f.

Gushee, Elizabeth M. 2008. "Travels through the Old South: Frances Benjamin Johnston and the Vernacular Architecture of Virginia." *Art Documentation: Bulletin of the Art Libraries Society of North America* 27 (1): 18–23.

Mitchell, Erik T. 2013. *Library Linked Data: Research and Adoption*. Library Technology Reports, Volume 49, Number 5. Chicago, Illinois: ALA TechSource.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

National Information Standards Organization (NISO). 2007. *A Framework of Guidance for Building Good Digital Collections*. Baltimore: www.niso.org/publications/rp/framework3.pdf.

Park, Jung-ran, and Yuji Tosaka. 2010. "Metadata Creation Practices in Digital Repositories and Collections: Schemata, Selection Criteria, and Interoperability." *Information Technology and Libraries*, 9 (3): 104–116. doi:10.6017/ital.v29i3.3136.

Shreeves, Sarah L., Jenn Riley, and Liz Milewicz. 2006. "Moving Towards Shareable Metadata." *First Monday* 11 (8). http://firstmonday.org/issues/issue11_8/shreeves/index.html.

Tani, Alice, Leonardo Candela, and Donatella Castelli. 2013. "Dealing with Metadata Quality: The Legacy of Digital Library Efforts." *Information Processing and Management* 49 (6): 1194–1205. doi:10.1016/j.ipm.2013.05.003.

Taylor, Arlene G., and Daniel N. Joudrey. 2004. *The Organization of Information*. Westport, CT: Libraries Unlimited.

University of Virginia Library Digital Initiatives. 2011. *General Descriptive Modeling Scheme (GDMS)* 9 (26). https://web.archive.org/web/20110926125213/http://www.lib.virginia.edu/digital/metadata/gdms.html.

Yasser, Chuttur M. 2011. "An Analysis of Problems in Metadata Records." *Journal of Library Metadata* 11 (2): 51–62. doi:10.1080/19386389.2011.570654.

Zeng, Marcia Lei, and Lois Mai Chan. 2006. "Metadata Interoperability and Standardization – A Study of Methodology Part II." *D-Lib Magazine* 12 (6). doi:10.1045/june2006-zeng.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

Susan Ivey and Michelle Emanuel

16 Large Scale with a Small Staff and Even Smaller Budget: Updating Metadata to Reflect Revised Best Practices

Introduction

The University of Mississippi Libraries (UML), in an effort to improve the metadata for its digital collections within aggregated environments, and specifically to bring it in line with the best practices of the Mississippi Digital Library (MDL) and the Digital Public Library of America (DPLA), embarked on a massive cleanup project of its existing collections in CONTENTdm.

Though UML's digital collections program was nearly ten years old, a best-practices document had never been finalized. This, coupled with a number of personnel changes, meant that metadata was ingested into CONTENTdm without a clear vision of how the collections might be used, or how the department might grow. Because the digital initiatives librarian position is part of the Department of Archives and Special Collections, and not housed within a department focused solely on digital initiatives, the position has been a "one man band" of sorts, requiring that the librarian build a close relationship with both Cataloging and Library Information Technology (LIT) departments. The issues of metadata best practices and of retrospective metadata cleanup were considered important, but were often not prioritized on the project schedule due to other, more pressing issues, such as (1) equipment purchasing, (2) digitization training of student workers, (3) finding aid markup, and (4) committee work. The application by the MDL, of which UML is a member, to join the DPLA pushed the retrospective project up on the project list.

A small team was assembled, comprised of team leaders (the Digital Initiatives Librarian and the head of Cataloging and Metadata Services), and three additional faculty members within the Cataloging and Metadata Services (chosen by the unit head and the Assistant Dean for Technical Services and Automation), in order to find creative solutions to spelling and format inconsistencies in the metadata that were very seriously limiting patron discoverability of and access to digital objects. The team recognized that existing digital collections were not intuitively named and did not use controlled vocabulary or consistent data formation. Moreover, much of the metadata was not created with the intention of being used outside of the local digital environment, and therefore the

https://doi.org/10.1515/9783110337419-017

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

use of field names and values needed to be normalized and enhanced in order to make metadata clearer and more usable to end users in an aggregated environment. These issues had led to collection-level issues, including field mapping to qualified Dublin Core and the collection level permissions and configurations, which was adversely affecting the discoverability of the digital collections. Lastly, the team agreed that future implementation of linked data should be considered during this process.

The project leaders, which include the Digital Initiatives Librarian and the Head of Cataloging, began by creating a metadata guidelines document based on best practices for participants in the MDL and the DPLA. The curators from Archives and Special Collections were trained on the new metadata practices, and the cataloging department began reviewing metadata for new projects prior to ingest, using tools such as OpenRefine. OpenRefine, which can be downloaded onto a local machine or used in a web interface, is an openly available application used to clean and transform data, and is particularly useful for large data sets. Once a new workflow was in place for new metadata creation, a plan to address legacy metadata housed within CONTENTdm based on the new policies had to be devised. The system made it easy to export the metadata, but impossible for it to be re-ingested onto existing objects. A review of the literature to understand how other institutions using CONTENTdm were addressing this issue uncovered very few examples. The project leaders were faced with evaluating and experimenting with workflows and solutions in order to determine the best method in which to proceed.

Background

In 2005, the Department of Archives and Special Collections within UML officially created the Digital Initiatives Librarian position with the primary goals of (1) encoding finding aids and (2) centralizing digitization projects and digital collections. In 2006, UML implemented CONTENTdm to make its digitized collections available online in a consistent format; however, the priority of the department remained on creating finding aids. From 2006–2009, digitization was occurring, yet with very little consistency in regards to imaging and metadata standards. Moreover, the collections that were added to CONTENTdm and those identified to become online digital exhibits as static Hypertext Markup Language (HTML) were chosen arbitrarily. This resulted in collections scattered throughout various platforms, which were also plagued with complications for browsing and searching due to inconsistent metadata practices. By 2009, thir-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

teen UML collections existed in CONTENTdm, and the department was beginning its first collaborative digital project as a part of the Association of Southeastern Libraries' (ASERL) *Civil War in the American South* project. ASERL specified required metadata fields, which UML followed. This is the first instance that documented metadata standards were imposed in the department.

In 2013, the current Digital Initiatives Librarian was hired. Because this individual is housed within the Department of Archives and Special Collections, there are many responsibilities incumbent upon this individual, such as (1) collaborating on exhibits, (2) providing instruction for visiting groups, (3) encoding finding aids, (4) acting on various committees, (5) serving the reference desk weekly, and (6) fulfilling tenure responsibilities. In terms of tasks associated with digital collections, the Digital Initiatives Librarian must (1) recommend equipment and software purchases, and manage their maintenance, operation, and training, (2) create and maintain collections within CONTENTdm, including mapping to Dublin Core (DC), (3) review and coordinate the ingestion of metadata and digital objects, and (4) collaborate with Library and Information Technology (LIT) department to troubleshoot errors that periodically occur on the locally-hosted CONTENTdm servers.

Of particular focus for the Digital Initiatives Librarian within the last two years has been (1) metadata standardization and remediation, (2) consistent file naming and file inventories, (3) creating and implementing a digital collection development plan, and (4) confronting digital preservation concerns. It was recognized during the spring of 2015 that these goals would require a large amount of time and resources. As such, the Digital Initiatives Librarian began collaborating with the Head of the Department of Cataloging, which had been recently renamed the Department of Cataloging and Metadata Services in order to create a streamlined workflow for metadata projects. The expected outcomes of this new cross-departmental collaboration were to create a team, comprised of the team leaders (the Digital Initiatives Librarian and the head of Cataloging and Metadata Services), and three additional faculty members within Cataloging and Metadata Services (chosen by the unit head and the Assistant Dean for Technical Services and Automation), in order to develop (1) a temporary, yet long-term project to overhaul legacy metadata within CONTENTdm and make it consistent with the newly created UML standards, and (2) a workflow for the creation of new metadata to increase authority control and to produce a larger amount of metadata in a shorter amount of time.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

Literature review

One vital step for this project was to create metadata guidelines that included a data dictionary for increased consistency for future collections, and would inform the overhaul of legacy metadata. While there was a vast amount of resources available to inform metadata creation and formation, the project team felt confident that basing their documentation largely on the best practices specific to the state's digital library, Mississippi Digital Library (MDL), would prove most beneficial. This is because the Digital Initiatives Librarian worked very closely with the MDL's Coordinator when drafting the Digital Public Library of America (DPLA) application. As such, the team knew these standards were up-to-date and best suited for aggregated environments. Recently, the digital initiatives librarian had been active on a metadata subcommittee for an aggregated project through the Association of Southeastern Research Libraries (ASERL) and the DPLA, and therefore felt confident that the knowledge gained from that project would mirror the current trends within the metadata field. Thus, the team decided that a literature survey on other areas of concern for this project would be the most beneficial as a next step. These areas included (1) the implementation of new cross-departmental metadata workflows, (2) examples of metadata projects involving CONTENTdm, and (3) case studies from libraries that were a part of the DPLA or that had applied or were considering an application.

Workflows

The literature regarding new metadata workflows among departments offered practical case studies that illustrated common strategies and challenges that could provide insight as UML implemented a newly formed cross-departmental metadata workflow. One example comes from the University of Montana's integration of metadata creation into the workflow of the traditional cataloging department, which was a direct result of participation in a nation-wide collaborative project Natives of Montana Archival Project (NOMAP) (Keenan 2012). A key lesson from the NOMAP project was the importance for both traditional catalogers and managers to be "flexible with priorities and reduce productivity in one area in order to increase it in another" (Keenan 2012, 208). Additionally, teaching new skills required for metadata creation should not and cannot be a one-time training session; managers should expect that some employees would require more training than others, and it is essential to offer a variety of training opportunities (Keenan 2012). Keenan (2012) describes the shifts in workflows as

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

an ever-evolving process, emphasizing that minor modifications and adjustments continue to be made in an effort to produce the best results.

Based on the lessons learned from the University of Houston's (UH) Digital Library pilot metadata audit project, Westbrook, Johnson, Carter, and Lockwood (2012) suggest that libraries rely on student intern help at every stage of a project whenever possible, because metadata projects require so much time to complete. The authors expressed surprise at the large amount of time it took to research and create their project plan, remarking that it took time away from the actual project, and therefore warn other institutions to be aware of the potentially daunting tasks associated with designing an audit project (Westbrook et al. 2012). The authors cite the reviewing of literature and drafting of the project plan as particularly time consuming due to the lack of examples available at the time of their project. However, the authors also note that as institutions conduct these projects and publish case studies there will be more of a foundation on which to design similar projects (Westbrook et al. 2012). Moreover, the authors strongly suggest that institutions make sure to create "robust record-keeping mechanisms early in the project" (Westbrook et al. 2012, under "Lessons Learned") as a means to keep the large amount of stakeholders informed throughout these often multi-year long projects, specifically citing shared collaborative cloud-based systems for remote work. Santi and Wu (2013) indicate that the pilot audit project led to an additional phase at the University of Houston Digital Library that outlined a new workflow, which included implementing (1) tab-delimited text files, (2) OpenRefine for increased quality control, and (3) an internal wiki to manage project workflows.

Khoo and Hall's (2013) case study about the creation and use of new metadata formats for Internet Public Library (ipl2) provides a strong contextual background from the literature about the complexity of metadata work due to (1) organizational gaps, including lack of professional training and expertise by those implementing metadata procedures, (2) limited staffing and time, (3) the complexity of metadata tools, and (4) the ever-evolving metadata standards. Khoo and Hall (2013) concluded that the ipl2 metadata study highlighted the "stickiness" of the knowledge required to work on metadata projects existing between networks of practice and across time, and that "legacy issues proved significant, and intercommunal negotiation had to be carried out across time, thought it was difficult to achieve" (102).

Phillips, Edward, Tarver, and Frakes (2014) describe the workflow developed for a collaborative digital project between the University of North Texas and the Oklahoma Historical Society. The authors emphasize that the flexibility of web-based metadata tools (specifically citing OpenRefine and Google Fusion Tables) is a valuable asset for collaborative metadata creation and mainte-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

nance. These tools drastically reduce manual editing of metadata post-ingest, while also allowing easy collaboration on one single document, making the division of work easier (Phillips et al. 2014).

Metadata workflow gaps based on "generational diversity" are also mentioned within the literature (Thompson 2015, 1). Thompson (2015) addresses the importance of communication, documentation, and an understanding of technology anxiety and change anxiety when "incorporating 'traditional' library staff into technology-rich workflows" (6). Of particular interest is Thompson's (2015) observation that catalogers, in particular, exhibited fear or anxiety about learning technological skills, and that digital staff had "long-standing internal perceptions" (3) that catalogers lacked the ability to learn these required skills. Therefore, metadata project-managers should be diligent in acting as ambassadors for the cataloging staff when bridging gaps among these groups (Thompson 2015).

CONTENTdm

A limited amount of sources exist about metadata upgrade projects specifically within the Online Computer Library Center (OCLC) CONTENTdm platform, which is surprising, due to the large amount of institutions that appear to be using CONTENTdm. According to results from a community survey and user interviews conducted by members of the Hydra-in-a-Box team, a project created by DPLA, Stanford, DuraSpace, and funded by the Institute of Museum and Library Services (IMLS), CONTENTdm was the most cited repository system from respondents (Hydra-in-a-Box 2015). Moreover, the team's competitive analysis of several repository systems also found CONTENTdm and Fedora to be the most widely adopted installations (Hydra-in-a-Box 2016). Therefore, due to the fact that the United States' largest aggregator, DPLA, which includes over a thousand donating institutions (DPLA Strategic Plan 2013) reported these usage numbers and results about CONTENTdm, it can be reasoned that there are a large number of institutions still using CONTENTdm. At the time of the team's literature scan, three published texts describing two institutions' experiences with CONTENTdm and legacy metadata upgrade processes were found.

Thompson and Wu (2013) outlined the University of Houston's Digital Library multi-year audit process, noting that they benchmarked their metadata and guidelines with comparable institutions' digital libraries that also used CONTENTdm. However, while this particular article supplied a detailed step-by-step process of the metadata audit project phases, it did not describe how the

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

upgraded legacy metadata was re-ingested into the system. A later publication with an additional author, however, did (Weidner, Wu, and Thompson 2014). The authors note that while exporting existing metadata from CONTENTdm and cleaning with OpenRefine was considered, the difficulties of re-ingesting the metadata to existing objects proved too difficult. Therefore, they chose instead to write applications in AutoHotkey to enable "efficient transformation of legacy authority data within the CONTENTdm Project Client" (Weidner, Wu, and Thompson 2014, 168).

The University of Utah's J. Willard Marriott Library chose to outsource their metadata enhancement to Backstage Library Works, with the goal to clean legacy data and to incorporate linked data Uniform Resource Identifier's (URI) (Myntti and Neatrour 2015, 195). The eXtensible Markup Language (XML) metadata was exported from CONTENTdm, Backstage ran authority control processes, and the XML metadata file was replaced on the CONTENTdm server (Myntti and Neatrour 2015). Errors that could not be fixed during the authority control process were corrected manually by an intern (Myntti and Neatrour 2015).

Digital Public Library of America (DPLA)

Since the University of Mississippi Libraries (UML) applied to be a part of DPLA, the team also scanned the literature for any insight about the use of metadata with DPLA, including (1) case studies from institutions that have become a part of DPLA, (2) those currently awaiting response to an application, and (3) those who are considering applying. In addition, the team researched any publications regarding metadata from the DPLA itself. Becoming aware of any metadata requirements or suggestions by DPLA or its participants prior to the implementation of the metadata project would ensure that the team did not have to duplicate or overwrite recent work providing the current application was accepted in the near future.

As a national digital library, many look to DPLA as a frontrunner in creating best practices for an aggregated digital environment. Hubs are requesting guidance from DPLA, and "while no formal collaboration has yet been established, DPLA now finds itself providing an important service, mediating connections across hubs to identify when the community faces common challenges" (Matienzo and Rudersdorf 2014, 8).

Staff at DPLA work closely with the hubs to identify quality issues at the time of acceptance. According to a report published by DPLA staff, after its first

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

year in operation, the hubs were "responsive and often eager" to make improvements, often creating tools to enrich metadata prior to ingest (Matienzo and Rudersdorf 2014, 7). At the time of this writing, though it is not a requirement that hubs rectify metadata to any particular standards prior to harvest, DPLA reported that it sees "promise in hubs taking on greater responsibility for metadata remediation, enrichment, and transformation to the [DPLA's metadata application profile] at the local level whenever possible" (Matienzo and Rudersdorf 2014, 9). The hubs also stated that they would like the DPLA to be able to share the metadata back to them in order to update their local metadata files, which DPLA described as one of its greatest challenges (Matienzo and Rudersdorf 2014). Because of these challenges, DPLA advocates for hubs to perform their own data remediation prior to sending it to DPLA, in order to promote "sustainability, improved data at the source, and speed of harvest," among other benefits (Rudersdorf 2014, 24). DPLA also encourages inputting data into more than their minimum required fields because "its user interface exposes and leverages fields like date, subject, and geographic coverage," and will expose these fuller records through its [Application Programming Interface] API (Gregory and Williams 2014, under "Choosing a Metadata Schema and Required Fields"). Because of these features, and in order to provide the cleanest metadata for greater interoperability, it appears that many individual institutions use their acceptance into DPLA as a reason to undergo legacy metadata projects (Gilbert and Mobley 2013; Gregory and Williams 2014; Matienzo and Rudersdorf 2014).

In terms of Service Hubs however, which are aggregators from multiple sources that act as the single point of contact for DPLA, it is unclear how many have altered metadata requirements for their members upon acceptance into DPLA. The North Carolina Digital Heritage Center chose to keep requirements from their member's low, in order to encourage participation from all types of institutions, recognizing that many may not have the skills or resources to undergo a massive metadata overhaul (Gregory and Williams 2014). While they note that their decision for requiring few fields to encourage participation was a good decision, it remains unclear how interested their data providers will be in improving their metadata in the future (Gregory and Williams 2014).

Pilot sampling phase

Before envisioning a workflow or creating a data dictionary and mapping schema, the project leaders and the team had to first investigate a sampling of existing digital metadata. This can be described as the pilot sampling phase, during

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

which team members would get an understanding of the types of inconsistencies that exist, and the project leaders would use trial and error to help inform the project design accordingly. This also gave the catalogers an opportunity to familiarize themselves with the CONTENTdm project client and with the software OpenRefine, which the project leaders selected for metadata cleanup. The digital initiatives librarian had experience with OpenRefine from past projects in which she participated at other university libraries (at that time known as GoogleRefine), and the literature also indicated that many metadata projects within academic libraries were currently using the tool. Additionally, the web services librarian at UML had advanced skills with OpenRefine, and volunteered to lead two workshops for the project team in order to give an overview of the tools' features. Therefore, due to its popularity within digital initiatives at academic libraries, and the fact that the tool is open source and therefore freely available to download on multiple machines, the project leaders chose OpenRefine as the first tool with which to experiment, confident that the tool would fulfill our requirements.

Since the digital initiatives librarian had worked closely with UML's CONTENTdm, she was aware of collections that had obvious metadata issues, and chose a variety of these with which the team members could experiment. These collections had a broad sampling of metadata issues, but less than 200–300 items, so as not to overwhelm the team.

After the training by the web services librarian, the Library Information Technology (LIT) Department installed OpenRefine on each member's computer and an exported tab-delimited text metadata file from CONTENTdm was sent to each team member. The files were opened in OpenRefine, and the data was examined using the clustering and faceting features, with detailed notes taken about the errors they found. This pilot sampling phase was mainly focused on the easily recognizable formation errors, such as inconsistent formatting of data, disregarding the actual meaning of the content in each field. Deeper content errors were examined in the active project phase. The project leaders and the team agreed that a three-week timeframe for this phase of experimenting would be sufficient.

When the project team met at the end of the sampling phase, each noted a large amount of errors within the metadata records. These errors include the typical errors seen within digital collection metadata, such as (1) misspellings; (2) inconsistent use of thesauri and locally created controlled vocabularies; (3) varied use of punctuation; and (4) trailing and leading white spaces.

While the catalogers continued to experiment with OpenRefine in order to become comfortable with its more advanced features (such as reconciliation support), the project leaders created a data dictionary that included mapping

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

specifications. This best-practices document would be used during the legacy metadata cleanup project, and would act as the guideline for the archivists, the catalogers, and the digital initiatives librarian when creating and reviewing new metadata for future collections. To create this guideline, the project leaders relied heavily on the guidelines created by the Mississippi Digital Library (MDL). The MDL's guidelines were originally created when the Mississippi Digital Library came into existence in 2003 and originally drafted by a team of catalog librarians from the University of Southern Mississippi. Changes to the guidelines occur periodically in a collaborative process that includes the catalog librarians in Bibliographic Services and Digital Collections at the University of Southern Mississippi. The MDL coordinator shared the mapping fields used by MDL, and the University of Mississippi Libraries (UML) and based its mapping specification on these.

Using the MDL's documentation as a foundation, the project leaders reviewed each field carefully, making changes that made more sense to UML's collections, such as requiring local controlled vocabularies. These included (1) type, (2) publisher, (3) contributor, (4) preferred citation, (5) original collection, (6) digital collection, (7) finding aid, and (8) capture method. During the active project phase, however, these fields were altered slightly, due to the legacy data that existed within the CONTENTdm shared controlled vocabulary lists, which were exported during the active project phase. The project leaders also added additional thesauri from those recommended in the Mississippi Digital Library guidelines such as the Art & Architecture Thesaurus and the Internet Assigned Numbers Authority (IANA) media types. Lastly, the project leaders chose to exclude a total of thirteen fields from the MDL guidelines, expecting these would not be relevant for UML's materials. This resulted in the UML data dictionary including thirty possible fields, with eight required.

Active project phase

Previous attempts to bring catalogers into the metadata workflow were stymied by difficulties with older versions of the CONTENTdm client. Many issues with permissions and sharing were resolved simply by having the project leaders use the web-based administrative module, while the other team members worked using Microsoft Excel. This change made the editing process much more efficient than working in the CONTENTdm project client, especially as the Excel spreadsheets can be loaded into OpenRefine and "cleaned" before they are ingested into CONTENTdm. It also acted as a control measure, ensuring that there

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

were not too many users with full administrative rights in CONTENTdm, thus eliminating the possibility of making permanent changes that could not be undone. Therefore, since the project leaders were the only members with administrative access to CONTENTdm, it became clear that they would be responsible for changes in mapping, updating controlled vocabulary, and collection level metadata edits, while the other members would oversee changes to legacy metadata.

Cleaning legacy metadata was problematic because in addition to editing the records themselves, it was necessary to clean the shared vocabularies, which included (1) personal and corporate names, (2) subject headings, (3) formats, and (4) collection names. The team learned early in the process that the controlled vocabulary lists in CONTENTdm had to be reviewed and altered prior to migrating the existing metadata. To deal with this task, the Head of Cataloging and Metadata Services opened the controlled vocabularies lists in OpenRefine to (1) review what data had been used within the collections, (2) make decisions regarding which entries to edit, and (3) to compare it to what was established in the best-practices documentation. CONTENTdm allows metadata to be approved even if a field contains values that are not included in the existing controlled vocabulary lists, and adds those values to the list after index and approval. While this feature can be helpful, it also can result in many inconsistencies if the administrator is not reviewing the vocabularies closely. As a result, in UML's case, these lists had become large and messy over time.

A variety of issues with these lists were uncovered. Many entries, for example, had been entered inconsistently resulting in various spellings and formatting of the same term. Moreover, some entries were incomplete. For example, several entries for names consisted simply of a first or last name. Lastly, many of the controlled vocabulary lists were combined, therefore containing multiple kinds of data. The "creator-contributor-subject" list for example existed as one large list. Separating these lists into individual field controlled vocabularies made it easier to review, and also makes maintenance of the lists easier in the future.

Because guidelines were never implemented, a number of collections included data that was unique to that collection in fields that would be difficult to search within aggregated environments. For example, in a collection of letters pertaining to the 1962 integration of the university, a field indicating whether the tone of the letter was "pro" or "con" was mapped to "original format." While a patron might be interested in such information, this information should have been included in a more appropriate field. Meanwhile, when cleaning the controlled vocabulary for original format, catalogers wondered to what "pro"

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

and "con" referred. These were valuable learning experiences, but were unexpected roadblocks in the project itself.

In terms of the individual object metadata, the project leaders decided to work on a per-collection basis in order to divide the work easily among the members, and to ensure that all records were reviewed at the same time, in order to view all possible inconsistencies. At this time, UML has ninety digital collections, varying in size from under ten objects to over 15,000 objects. Project leaders decided that for collections with fewer than one hundred items, they would (1) export the metadata, (2) clean it in OpenRefine, and (3) paste it back in the project client. However, for collections exceeding one hundred items, they looked for a way to clean and reload the metadata. One option included deleting all contents from the collections, and re-ingesting the objects and the metadata after the metadata was reviewed in OpenRefine, although they acknowledged this may not be the in keeping with best practices. They also wondered how easy it would be to find the location of the objects on the archival server, since inconsistency in file naming and file storing conventions at UML was another issue being faced. However, they realized that since the projects were locally hosted, they could use the existing files on the CONTENTdm servers to re-ingest. To do this, they would need to locate and copy the files in the "image" subfolder of the collection folder to a new location, and use the CONTENTdm assigned an Online Computer Library Center (OCLC) filename from the exported metadata as the new object filename. At the time of this writing, they were planning to try a smaller test collection utilizing this method. If this method proved successful, they would do this for all collections containing over one hundred items.

The digital initiatives librarian would oversee the collection-level issues, including field mapping to qualified Dublin Core and the collection level permissions and configurations. This would involve going through each collection to add or delete fields according to the data dictionary, and to ensure proper mapping. This would be done during the post-processing phase once it were clear which fields would be used in each collection.

Conclusion

While the project leaders expected the project to be time-consuming and recognized the need to remain flexible in order to problem solve quickly and effectively, the amount of obstacles to overcome and the time it took to address these were still unexpected. The most complex task included creating the project de-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

sign, as case studies in the literature indicated. Since they lacked step-by-step examples dealing with legacy metadata within CONTENTdm, they found themselves learning by trial and error, figuring out the order required for changes within the system. With what is expected to be a growing number of institutions addressing legacy metadata within CONTENTdm, they hope that more detailed examples appear in the literature, particularly for the benefit of smaller libraries lacking programming and other technical resources.

The project team had also begun experimenting with the reconciliation feature in OpenRefine. While they were not ready to begin implementing linked data practices, they recognized its importance and wanted to use this project as a means to begin understanding how OpenRefine might help them with this transition in the future. They also hoped that more specific examples from the library community would emerge regarding this point.

Lastly, they understood that metadata requires continuous maintenance, and therefore recognized the need to create a process to regularly review and update metadata and its associated guidelines. Moreover, as aggregated collections continue to grow, and as they continue to participate in projects on a larger scale, they need to think about their metadata in a broader context, challenging themselves even further to consider future uses of the data outside of the current environments in which they are now sharing.

For institutions like the University of Mississippi Libraries, without a single designated department devoted to digital initiatives, it is vital that all departments involved have an open and honest dialogue about what would be required to accomplish the institution's goal. While the team might wish that different decisions had been made at the beginning of the digital program, they could take advantage of the innovations adopted in the meantime, such as OpenRefine, to get the program back on track.

References

Digital Public Library of America (DPLA). 2013. "DPLA Strategic Plan 2015–2017." https://dp.la/info/wp-content/uploads/2015/01/DPLA-StrategicPlan_2015–2017-Jan7.pdf.

Gilbert, Heather, and Tyler Mobley. 2013. "Breaking up with CONTENTdm: Why and How one Institution Took the Leap to Open Source." *The Code{4}Lib Journal* 20. http://journal.code4lib.org/articles/8327.

Gregory, Lisa, and Stephanie Williams. 2014. "On being a Hub: Some Details behind Providing Metadata for the Digital Public Library of America." *D-Lib Magazine* 20 (7/8). doi:10.1045/july2014-gregory.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

Hydra-in-a-Box. 2015. "User Interviews and Focus Groups Documentation." https://wiki.duraspace.org/display/samvera/Hydra-in-a-Box+Focus+Group+Lunch+at+Open+Repositories+2015

Hydra-in-a-Box. 2016. "Competitive Analysis Documentation." https://wiki.duraspace.org/display/hydra/Hydra-in-a-Box+Design+Documents?preview=%2F76841397%2F77431612%2FHydra-in-a-Box+Competitive+Analysis.pdf

Keenan, Teressa M. 2012. "Charting a Course with NOMAP: Integrating Metadata Workflows into a Traditional Cataloging Unit." *Journal of Library Metadata* 12 (2–3): 199–212. doi:10.1080/19386389.2012.699842.

Khoo, Michael, and Catherine Hall. 2013. "Managing Metadata: Networks of Practice, Technological Frames, and Metadata Work in a Digital Library." *Information and Organization* 23 (2): 81–106. doi:10.1016/j.infoandorg.2013.01.003.

Matienzo, Mark A., and Amy Rudersdorf. 2014. "The Digital Public Library of America Ingestion Ecosystem: Lessons Learned After One Year of Large-Scale Collaborative Metadata Aggregation." In *International Conference on Dublin Core and Metadata Applications*. http://dcpapers.dublincore.org/pubs/article/view/3700/1923.

Myntti, Jeremy, and Anna Neatrour. 2015. "Use Existing Data First: Reconcile Metadata before Creating New Controlled Vocabularies." *Journal of Library Metadata* 15 (3–4): 191–207. doi:10.1080/19386389.2015.1099989.

Phillips, Mark Edward, Hannah Tarver, and Stacy Frakes. 2014. "Implementing a Collaborative Workflow for Metadata Analysis, Quality Improvement, and Mapping." *The Code{4}Lib Journal* 23, http://journal.code4lib.org/articles/9199/comment-page-1.

Rudersdorf, Amy. 2014. "Implementing Collection Contexts, and Metadata Issues Related to Normalization and Shareability." In *Digital Collection Contexts: iConference 2014 workshop report*, by Palmer, Carole L., Antoine Isaac, Karen M. Wickett, Katrina S. Fenlon, and Megan Senseney, 24–26. Urbana-Champaign: Center for Informatics Research in Science and Scholarship.

Thompson, Kelly J. 2015. "'What If I Break It?': Project Management for Intergenerational Library Teams Creating Non-MARC Metadata." *The Code{4}Lib Journal* 28, http://journal.code4lib.org/articles/10395.

Thompson, Santi, and Annie Wu. 2013. "Metadata Overhaul: Upgrading metadata in the University of Houston Digital Library." *Journal of Digital Media Management* 2 (2): 137–147.

Weidner, Andrew, Annie Wu, and Santi Thompson. 2014. "Automated Enhancement of Controlled Vocabularies: Upgrading Legacy Metadata in CONTENTdm." In *Proceedings of the 2014 International Conference on Dublin Core and Metadata Applications*, 167–172. Dublin Core Metadata Initiative. http://dcevents.dublincore.org/IntConf/dc-2014/paper/view/218.

Westbrook, R. Niccole, Dan Johnson, Karen Carter, and Angela Lockwood. 2012. "Metadata Clean Sweep: A Digital Library Audit Project." *D-Lib Magazine* 18 (5/6). doi:10.1045/may2012-westbrook.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:09 AM

Samantha Norling

17 Bringing the Archives Out of the Art Museum: Local Metadata Planning within a Global Context

Introduction

In 2015, the staff of the Indianapolis Museum of Art (IMA) Archives was preparing to ingest digitized archival collection images and associated metadata into a new institution-wide digital asset management (DAM) system. The addition of this material to the DAM marked the first time that IMA Archives' collections images and metadata would be stored and discoverable within the same system as images and metadata related to the museum's permanent collection of objects and artwork. Given this opportunity for digital convergence, and with the ability to completely customize metadata fields, the archivists took a "think globally, act locally" (Waibel and Erway 2009, 323) approach to metadata planning. Looking past the immediate descriptive needs of the IMA Archives, staff took into consideration the descriptive practices for the permanent collection and looked more broadly at existing metadata documentation for both the state-wide digital library, Indiana Memory, and the national platform, the Digital Public Library of America (DPLA). By taking this approach, the IMA Archives ended up with a metadata schema that maximizes opportunities for digital convergence both locally, between the IMA's archival and art collections, and globally, by mapping to a national platform and prescribing widely-accepted controlled vocabularies, data formats, and accepted rights statements for relevant fields.

The global context

There is a long-standing history of collaboration within each of the library, archives, and museum (LAM) professions individually. This tradition is evident through the broad acceptance of practices such as (1) shared cataloging, (2) the adoption of standardized data structures such as MAchine-Readable Cataloging (MARC) records and Encoded Archival Description (EAD) to represent collection information online, and (3) the use of established controlled vocabularies and

https://doi.org/10.1515/9783110337419-018

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

thesauri in the description of special collections. These shared practices and standards form the backbone of union catalog initiatives such as WorldCat (for libraries) and ArchiveGrid (for archives). While these discovery portals are powerful, centralized resources for online users, they illustrate how shared practices and standards within the separate LAM professions have contributed to the isolation of cultural heritage institutions from one another (Waibel and Erway 2009). This history of collaborative professional practices and standards has been iterative, and the recent trend of LAMs embracing linked open data is the next logical step for libraries, archives, and museums navigating within the web 3.0 environment. What makes linked data unique within this tradition is its capacity for cross-disciplinary application, and the result has been an emerging awareness among LAM institutions and professionals of the need to break down traditional barriers to bring disparate collections together in the digital realm.

The emergence of web 3.0 and related technologies has placed an increasingly large emphasis on generating and enriching connections among information, people, and applications in the online environment. As systems, data structures, and standards have evolved to support the creation of these connections, the result has been an increase in collective intelligence generated by individuals, groups, and organizations that are actively creating, sharing, and connecting information on the web (Spivack 2007). For LAMs, this means that while physical collections remain necessarily separated by institutional boundaries, online users of these collections increasingly expect to discover and explore the full scope of available information from a single search. The recent emergence of digital platforms such as Europeana and the Digital Public Library of America has made it possible for LAMs of all types and sizes to not just display their digital and digitized collections online, but to more meaningfully connect this material within the broader scope of linked data on the web.

Launched to the public in 2013, the Digital Public Library of America (DPLA) is a rapidly growing initiative to "[bring] together the riches of America's libraries, archives, and museums, and [make] them freely available to the world" (DPLA 2016, "About"). To accomplish this vision of a unified discovery portal for digital cultural heritage collections, the DPLA has established a growing network of "Content Hubs" and "Service Hubs" (Figure 17.1) from which metadata and thumbnail images are regularly harvested for inclusion in DPLA. Within this model, Content Hubs are individual LAMs that contribute a large amount of digital content directly to the DPLA. In contrast, Service Hubs are "state, regional, or other collaborations that host, aggregate, or otherwise bring together digital objects" (DPLA 2016, "Hubs") from LAMs for delivery to DPLA. At the time of this writing, the DPLA hub network includes sixteen Content

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

Hubs and 21 Service Hubs, including Indiana Memory, the statewide Service Hub through which the IMA Archives' collections are aggregated.

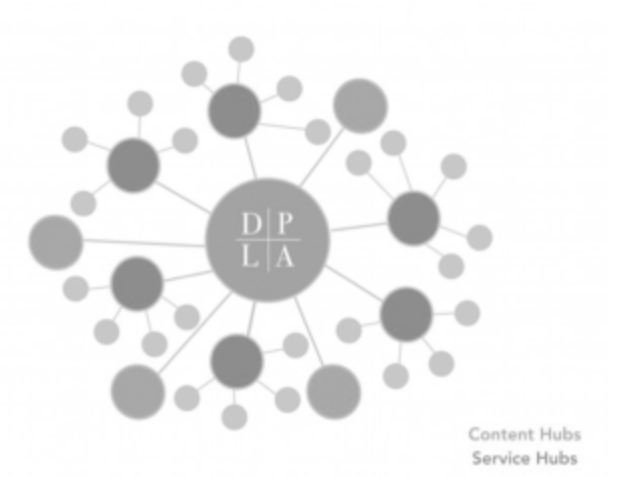

Fig. 17.1: Model of DPLA's network of Content and Service Hubs DPLA, Inc. / CC-BY-SA-3.0

With an organizational culture that emphasizes open access, a major component of the DPLA model is an application programming interface (API) that provides direct access to aggregated metadata in the form of JavaScript Object Notation for Linked Data (JSON-LD), a linked data format (Mitchell 2013). As a result, individual LAMs contributing to DPLA end up also contributing to the growing network of linked open data on the web, even if they are not actively encoding their metadata in a linked data format within their institution.

The local context

As linked open data continues to gain traction throughout the LAM community, institutions may find that their existing systems for managing collections data and digital assets no longer meet the needs of an increasingly connected online environment. The Indianapolis Museum of Art (IMA) found itself in just such a situation in 2012. At the time, IMA Photography Department assets lived in four separate places due to a rushed DAM implementation, multiple workflow changes during that implementation, and minimal integration between the DAM and the Museum's collection management system. The existing DAM was

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

also lacking some key features and utilities, including (1) no ability to run reports and download statistics about the collection, (2) no metadata field for capturing rights information or credit lines, (3) no batch metadata creation function, and (4) inconsistent syncing of metadata with images. As a result, processing time for new photography and rights clearance requests was slow and marketing and other departments' ability to access images and generate accurate captions was limited. In short, existing systems did not allow for maximum leveraging of the IMA's digital assets and collections data, to the detriment of both external users and Museum staff. This reality prompted the IMA to begin the multi-year task of researching and implementing a new DAM. At the end of this research and evaluation process, the task force recommended Piction as the system that would best meet the IMA's digital asset and metadata management needs.

Phase I of Piction implementation was carried out over a year, and involved multiple steps, such as the (1) preparation and clean-up of existing permanent collection assets and metadata, (2) mapping metadata fields from the CMS to fields in Piction, (3) the technical back-end integration of the two systems, (4) multiple bulk ingests of assets, and (5) the creation of various workflows for not only the ingest of assets and metadata, but also for processes such as rights clearance and new photography requests. Up to this point in the implementation process, the primary focus had been on the Museum's permanent collection of objects and artwork, with plans to add digital archival collections to Piction during Phase II. In order to fully understand the reason for this separation, and to provide additional context for decisions made while implementing Piction for the Archives, a bit of background information is needed about the role of the IMA Archives within the Museum, and of its relation to the permanent collection.

The archives in the art museum

The IMA is an encyclopedic art museum located on the north side of downtown Indianapolis, Indiana. The Museum's 152-acre campus includes the Virginia B. Fairbanks Art and Nature Park and the historic Oldfields estate of the Lilly House and Gardens. Since 2009, the IMA has also stewarded the Miller House and Garden property located in Columbus, Indiana. The main museum building in Indianapolis includes two theatres, event spaces, staff offices, and a library, along with the storage and gallery spaces where the permanent collection of more than 54,000 works representing 5,000 years of the world's cultural herit-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

age resides. While the IMA was founded over 130 years ago, the IMA Archives was not formally established until 2010, when the Museum hired its first archivist. As one half of the Library & Archives Department of the IMA, the archives as a collection serves two functions: first, as the primary repository for institutional records of all types, in both physical and digital formats; and second, as the home of archival collections documenting notable individuals, art communities, collectors, and artists in Indiana, many of which have direct ties to the history of the IMA and its collections. The primary users of the archives are IMA staff members, though many of the records and manuscript collections are open to external researchers.

While some of the archival collections are only peripherally related to the museum's collection, activities, and history, many of the largest record groups in the archives relate directly to the permanent collection. Most notable of these are records documenting hundreds of exhibitions organized by or held at the Indianapolis Museum of Art and its two predecessors, the Art Association of Indianapolis and the John Herron Art Museum. Nearly all of the exhibition documentation contains information of value to the Curators and Registrars in their role as permanent collection managers, such as (1) ephemeral brochures and catalogs, (2) incoming and outgoing loan receipts, (3) final exhibition checklists listing all items exhibited, (4) condition reports, (5) curatorial research into the provenance of artwork and the biographies of artists, (6) installation photographs, and (7) general correspondence which frequently discusses both considered and realized acquisitions to the Museum collection.

In the five-year history of the IMA Archives, the visibility and relevance of the archival collections, both within and outside of the institution, has steadily improved, owing to a series of important acquisitions and projects.

- 2009: The Miller House and Garden Collection (M003) of archival records documenting the design, construction, and maintenance of the iconic property came to the Museum along with the house and grounds. This acquisition served as a major impetus for hiring a professional archivist to preserve the collection and make it available to IMA staff in charge of historic preservation and interpretation planning of the property.
- 2010: The newly-hired archivist began creating detailed inventories of institutional records and special collections previously managed by Registration, and opening them up to researchers.
- 2011: Through a Library Services and Technology Act (LSTA) grant from the Indiana State Library, digitization of the Pioneer Painters of Indiana Research Collection (M002) was completed.
- 2012: The Archives received a grant from the National Endowment for the Humanities (NEH) to digitize the Miller House and Garden Collection.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

Through the funds and cost share, the IMA Archives significantly built up its capacity for in-house digitization, and more than tripled the number of digital assets and related metadata records managed by the archivist.

– 2014: Two major collections of art collectors' records were acquired and made available for research. One of the collectors, Onya La Tour, was the subject of an IMA exhibition in the art galleries, for which a curator selected archival material to display alongside the artwork.

With the improved visibility of the IMA Archives, the archivists' capacity to create and manage digital assets and information also grew, whether in the form of lengthy Encoded Archival Description (EAD) collection guides or digital surrogates of archival items with associated metadata. This growing emphasis on the creation of digital collection descriptions published to the web is in no way unique to the IMA Archives. It has been a major growth area within each LAM community for multiple decades, particularly in the area of digitization efforts. One result of this increased availability of and reliance on these online resources in the LAM community has been a "digital convergence of libraries, archives, and museums... [a blurring] of traditional distinctions between information organizations" in the digital realm (Marty 2009, 295). For the IMA Archives, this blurring of distinctions, particularly between that of archival and museum collections, is key to raising the relevance of the archival collections in relation to the permanent collection to both internal and external audiences.

Metadata planning for local digital convergence

Prior to the implementation of Piction for digital asset management across all departments at the IMA, the permanent collection and the IMA Archives collections did not have available avenues for digital convergence. Data about the permanent collection was stored in the Museum's CMS, KE-EMu, while archival descriptions and metadata were maintained in Archon, museum- and archives-specific software, respectively. Digital images and associated descriptions of the permanent collection (originating from the Photography, Curatorial, and Registration Departments) were managed and accessed primarily through a past DAM, while digital surrogates of physical archival items were accessible in Archon as digital objects with metadata linked to the EAD finding aids. Migrating both groups of digital assets and associated metadata into a unified DAM would allow IMA staff to discover relevant items from both collections from a

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

single search, an ideal that most people have come to expect in today's linked digital environment.

One of the most beneficial characteristics of Piction that made this DAM appealing to the IMA task force during their research was its flexibility in terms of metadata structure and content. While the system supports the use of multiple standards commonly used by LAMs, institutions also have the ability to customize metadata schemas at the collection level. For the IMA, this feature was crucial. Piction would not only need to accommodate the differing descriptive needs of both the permanent and archival collections, but also the needs of special events, marketing, and other non-collection digital assets. Presented with this blank slate for metadata planning, the archivists were eager to develop a metadata schema which would not only further internal digital convergence with the permanent collection, but which would also facilitate easier mapping to the requirements of DPLA, to which the archivists hoped to one day contribute. For the metadata planning process, the archivists followed the advice of Waibel and Erway (2009), that "while it is important to be mindful of this global network vision, it is equally important to focus on the local actions which will allow this vision to come to fruition. It seems unlikely that LAMs can successfully collaborate on the network level unless they have worked through issues of convergence in their own backyard" (325). While keeping in mind the ultimate goal of contributing IMA Archives digital collections to statewide and national digital libraries, the focus would first be on finding opportunities to leverage descriptive archival metadata to establish meaningful connections with the IMA's permanent collection.

To this end, IMA Archives staff reviewed use statistics that had been collected since the founding of the archives in 2010. Looking at the topical subjects of reference questions received from IMA staff revealed that the top areas of inquiry were (in order of importance), (1) IMA exhibitions (27 percent of all staff requests, 2010–2015), (2) Miller House and Garden (15 percent), (3) donors/development (12 percent), (4) permanent collection items (seven percent), and (5) artists represented within the permanent collection (five percent). The high level of staff interest in the records documenting the many exhibitions held by and/or at the IMA was not surprising. This collection is expansive and includes documentation originating from most departments within the Museum from the time of its founding to the present day. As previously discussed, it includes many document types that contain images and both contextual and physical descriptions of items in the IMA's permanent collection and artists represented in that collection. In this way, exhibition records are often a primary source used by staff researching the collection pieces, especially when establishing provenance and exhibition histories.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

With these research priorities in mind, IMA Archives staff then reviewed the descriptive practices for the permanent collection, specifically the metadata fields employed by the registrars and curators, to look for opportunities of convergence with archival description. While hundreds of individual data points are captured in the Museum's CMS to describe each art piece, less than two dozen fields are pulled from the CMS to then populate metadata fields in Piction. These fields represent a mix of descriptive and administrative metadata.

Many of the descriptive metadata fields are populated with values from prescribed controlled vocabulary or thesauri, such as the Artist/Designer field, which utilizes authorized headings from the Union List of Artist Names (ULAN). While many of these metadata fields, particularly the administrative pieces of information, do not have a direct relation to the archival collections, the archivists used this list to consider the possibilities of incorporating specific pieces of information, if not entire metadata fields, into the archival metadata schema. It was at this point in evaluation that the decision was made to utilize some of the same controlled vocabularies and thesauri as prescribed for the permanent collection metadata, where applicable, in the new Archives metadata schema.

Combining statistics about staff research interests with the metadata already provided in Piction about permanent collection assets, IMA archivists were forming a clearer picture of which pieces of information would be most valuable to capture in the archival metadata in order to meet the research needs of IMA staff and to further the internal digital convergence between the two collections. The short list included (1) IMA accession numbers (unique identifiers assigned to specific items in the permanent collection), (2) artist names (values prescribed by ULAN), (3) art movements/periods/styles (prescribed by the Art & Architecture Thesaurus [AAT]), (4) materials/medium (prescribed by AAT), (5) names of individuals who donated pieces to the Museum, and (6) titles of exhibitions held by and/or at the IMA.

Local metadata planning within a global context

Continuing the preparation for Phase II of Piction implementation, the next step that the archivists took was to gather information and resources that would help to build a knowledge base. It is important to note that metadata experts with years of experience in the field did not conduct the evaluation of existing metadata and all subsequent steps in creating a metadata schema for the IMA Archives. The responsibility for creating the schema fell to the IMA's archivist, who was one year into her first professional position and whose previous expe-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

rience was largely with physical archival collections and traditional descriptive methods. A dual library science/public history graduate student interning with the IMA Archives at the time, Rebecca Pattillo, assisted with the metadata planning. Recognizing their limitations, preliminary research and resource gathering was essential. Of particular note is the publication *Metadata for Digital Collections* by Steven J. Miller (2011), which served as a primary reference source throughout the project, along with metadata documentation for both DPLA and Indiana Memory, two digital libraries to which the IMA Archives could potentially contribute. It was during this information-gathering phase that the decision was made to base the new IMA Archives metadata schema primarily on Qualified Dublin Core (DC), as many of its elements were included in both digital libraries' metadata documentation.

After building up a knowledge base, the next step for IMA archivists was to evaluate the archives' existing metadata. At that time, there were only two collections from the IMA Archives that had been digitized on a large scale: the Pioneer Painters of Indiana Collection and the Miller House and Garden Collection (MHGC), which was still in the process of being digitized. The Pioneer Painters Collection had been digitized at a time when the IMA Archives had no capacity to carry out large-scale digitization in-house. For that reason, an outside vendor, the Indiana University-Purdue University Indianapolis (IUPUI) Center for Digital Scholarship, had been contracted to digitize the collection, with Library & Archives Department staff then creating related metadata in a basic Microsoft Excel spreadsheet. When completed, the metadata was ingested into CONTENTdm, which supports the creation of metadata in the Qualified Dublin Core schema. Though DC elements were not specifically listed on the metadata spreadsheets created for this collection, each field was mapped to one of the DC elements by IUPUI staff, and that mapping was available for referencing.

In order to gather standard information about the metadata for both archival collections, a Metadata Evaluation Template was created to collect the information outlined in Table 17.1.

Tab. 17.1: IMA Archives Metadata Evaluation Template Column Headers with Descriptions

Field Name	(Q)DC Element	Legacy Vocab/Thesauri	Preferred Vocab/Thesauri	Value Example and Notes
Title of existing metadata field	Corresponding Dublin Core element to map metadata field to	Controlled vocabularies and/or thesauri used in existing metadata	Preferred controlled vocabularies and/or thesauri, to reference during metadata clean-up	Example metadata value when needed, and any special notes regarding the field

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

After plugging the metadata fields from the Pioneer Painters of Indiana spreadsheet into the Metadata Evaluation Template, the process of selecting a corresponding DC element and adding relevant notes to each of the template fields was simple. Additional vocabularies preferred or required by DPLA and Indiana Memory were then added to the evaluation spreadsheet. During Phase II preparation, IUPUI was still hosting the Pioneer Painters Collection on their servers, so the IMA archivists were working with the metadata in the original Microsoft Excel spreadsheet format. This made clean-up and modification much simpler than if the data had to be mapped, pulled, and then migrated from one system to another, which was the case with the MHGC.

As previously mentioned, the metadata for MHGC had been created in Archon, where the records linked directly into the EAD collection guide. Before this data could be ingested into Piction, the IMA database administrator first pulled the data from Archon and made the more than 9,400 records available for the archivists in a comma separated value (CSV) format, which could be viewed as a spreadsheet and cleaned up in bulk prior to ingest. As with the Pioneer Painters collection, the metadata fields from MHGC data records were plugged into the evaluation spreadsheet to map to Qualified DC elements with notations about content values (Table 17.2).

Tab. 17.2: Miller House and Garden Collection (M003) Metadata Evaluation Spreadsheet

Archon Field Name	(Q)DC Element	Legacy Vocab/Thesauri	Preferred Vocab/Thesauri	Value Example and Notes
Title	Title			
Identifier	Identifier			[unique ID number for item(s) related to metadata record]
Date	Date.Original	EDTF; local	EDTF	[local formatting includes "ca.", "n.d."]
Languages	Language	Archon-specific code		[repeatable] Spell out in full: e.g. "English"
Subject (combined w/below)	Subject	LCSH; AAT; local	LCSH; LCNAF	[repeatable]
Subject (combined w/above)	Format	AAT	AAT	[repeatable]
Creator Information	Creator	LCNAF; ULAN; local	LCNAF; ULAN	[repeatable]
Scope	Description			Full transcription following "Typewritten text:"
Physical Description	Format			eg. "5 x 8 in."
Publisher	Publisher	LCNAF; local	LCNAF	
Rights Statement	Rights	local	DPLA Rights	[local: assertion of fair use]
Collection	RelationisPartOf			"Miller House and Garden Collection, 1953-2009"
Collection Content	RelationisPartOf			[physical container]
File Title	Description			[image filenames, repeatable]

A comparative analysis of the two collections indicate both had been described with significant variation in metadata. The goal of the archivists at this stage was to develop a unified IMA Archives metadata schema, mapped to Qualified DC, which these two collections could be adjusted to fit, and that would also serve the descriptive needs of collections that would be digitized in the future. This schema would also need to include fields to capture the specific pieces of information of high research interest to IMA staff. Although the IMA Archives

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

was still well over a year away from contributing digital collections to DPLA, it was decided that the new metadata schema should also include all fields required or preferred by DPLA, even if they had not been utilized in the existing collections, with specific notation of DPLA preferences incorporated into the schema. To that end, DPLA's Metadata Application Profile (MAP) became a key reference document in the final stage of creating the IMA Archives' new metadata schema for Piction.

Tab. 17.3: IMA Archives Simple Metadata Schema with DPLA Notations.

(Q)DC Element	Piction Field Name	DPLA Field	Required by DPLA?	Displayed in DPLA?	Vocabulary/Thesauri
class="dpla:SourceResource"					
Title	Title	Title	Yes	Yes	
Identifier	UUID		No	No	
Identifier	Identifier		No	No	
Date.Original	Date	Created Date	Yes	Yes	EDTF
Language	Language	Language	Yes	Yes	
Subject	Subject	Subject	Yes	Yes	LCSH; LCNAF; ULAN; AAT
Creator	Creator	Creator	Yes	Yes	LCNAF; ULAN
Description	Description	Description	No	Yes	
Description	Transcription		No	No	
Format	Physical Description	Format	Yes	No	
Publisher	Publisher	Publisher	Yes	Yes	LCNAF
Rights	Copyright	Rights Information	No	Yes	legacy; RightsStatements.org
RelationisPartOf	Collection	Collection	Yes	No	
RelationisPartOf	Finding Aid	Collection Description	No	No	
RelationisPartOf	Physical Container		No	No	
RelationReferences	Accession Number		No	No	
RelationReferences	Exhibition		No	No	[local exhibition vocabulary]
Type	Type	Type	Yes	Yes	DCMI Type Vocabulary
Format	Format	Format	Yes	No	AAT (form of material)
Description	Preferred Citation		No	No	
class="dpla:Place"					
Coverage (spatial)	Geographic Location	Place	Yes	Yes	LCSH
class="edm:WebResource"					
Description	Filenames		No	No	
edm:Rights	[only used for DPLA]	Standardized Rights	Yes	Yes	RightsStatements.org
class="ore:Aggregation"					
Contributor	[only used for DPLA]	Contributing Institution	Yes	Yes	"IMA Archives"
edm:IsShownAt	[only used for DPLA]	URL	Yes	Yes	[persistent URL]
edm:Preview	[only used for DPLA]		Yes	No	[thumbnail URL]

By combining all of the documentation created and gathered for reference throughout this process, the archivists found it surprisingly simple to bring together all of the pieces of information into one unified metadata schema (Table 17.3).

The DPLA MAP provided the framework for this new schema, which divides the metadata fields into four core classes adapted from the MAP: (1) SourceResource: attributes of item(s) being described ("Source Resource"), not the digital

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

surrogate, (2) Place: geographic information, (3) WebResource: attributes of the digital surrogate ("Web Resource"), (4) Aggregation: attributes of the digital surrogates in aggregate.

One notable change to the existing metadata appears in the Subject field. The Pioneer Painters Collection included six separate fields that mapped to dc: Subject, and both existing collections utilized the Subject field to capture information that was more appropriate for a Format field. In the unified schema, there is only one Subject field and one Format field. This required that the multiple subject values captured in the Pioneer Painters metadata be consolidated into a single Subject field that accommodates multiple values, and for the Format values to be parsed into a separate field. Bulk clean-up of these values was achieved through the use of OpenRefine.

Two fields in the unified metadata schema were added specifically to serve the research needs of IMA staff, Accession Number and Exhibition (both mapped to dc:RelationReferences). The Accession Number field is notable in that it does not capture the accession number for the archival resource being described, but is instead a place to document references made in the archival resource to artwork in the permanent collection. As the accession number serves as the main identifier for each permanent collection item, it is the preferred search term used by IMA staff in Piction, making this field a key addition to the schema. Similarly, the Exhibition field is where references made to IMA exhibitions are documented, utilizing a controlled vocabulary developed locally. Three additional fields, (1) Subject, (2) Creator, and (3) Transcription, are also highlighted in the schema documentation to indicate to metadata creators that they should be populated specifically with IMA staff research needs in mind, with the goal of establishing meaningful connections between the archival item being described and related items in the Museum's permanent collection, where applicable.

Conclusion: Local and global digital convergence

The result of this purposeful metadata planning has been twofold. Within the Museum, the IMA Archives has evolved from a separate, supporting collection of historical records to reference when needed into a vital collection of records playing a more active role in current interpretation and collecting activities. Designing a metadata schema with staff needs in mind has furthered this evolution along at a pace that would not have been possible had digitized archival material been added to Piction without the descriptive metadata specifically

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

tailored to connect with the permanent collection. Not only has the metadata planning resulted in a heightened relevance and awareness of the Archives within the IMA, but it has also contributed significantly to growth in the Archives' visibility outside of the Museum. While the archivists were focused on internal benefits when incorporating museum collection-specific metadata fields into the schema, the information that those fields capture is now benefiting external audiences through the IMA's Digital Archives Portal (http://archive.imamuseum.org). Launched in 2015, the Portal is the online discovery interface for the digital archives collections, pulling images and metadata from Piction.

By making the collections available online, the launch of the Digital Archives Portal made it possible for the archivists to begin the process of contributing to DPLA. In June 2016, the IMA Archives became the first institution that was not using CONTENTdm to contribute to Indiana Memory to stage its collections for DPLA harvesting. The selection of the IMA Archives to serve as the test case for the new DPLA Indiana Metadata Mapping Tool owed almost exclusively to the fact that the archivists had already worked through mapping their metadata fields to DPLA. As a test case, the IMA Archives metadata mapping was a great success, aided not only by the previous mapping, but also by the fact that the data complied with many controlled vocabularies and formats required or preferred by Indiana Memory and DPLA. In this case, intentional metadata planning brought the archives out of the art museum by making it possible for its collections data to join the growing network of linked information on the web, results that many LAM institutions with all types of cultural heritage collections could realize by adapting the "think globally, act locally" approach to their digital descriptive practices.

References

Digital Public Library of America (DPLA). 2016. "About." https://dp.la/info/.

Digital Public Library of America (DPLA). 2016. "Hubs." https://dp.la/info/hubs.

Indiana State Library. 2007. "Dublin Core Metadata Guide: Indiana Memory Project." http://www.in.gov/library/files/dig_metast.pdf.

Marty, Paul F. 2009. "An Introduction to Digital Convergence: Libraries, Archives, and Museums in the Information Age." *Museum Management and Curatorship* 24 (4): 295–298. doi: 10.1080/09647770903314688.

Miller, Steven J. 2011. *Metadata for Digital Collections: A How-To-Do-It Manual*. How-To-Do-It Manuals, No. 179. New York: Neal-Schuman.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

Mitchell, Erik T. 2013. *Library Linked Data: Research and Adoption*. Library Technology Reports, Volume 49, Number 5. Chicago, Illinois: ALA TechSource.

Spivack, Nova. 2007. "Enriching the Connections of the Web—Making the Web Smarter." *Minding the*

Planet (blog), July 3. http://www.novaspivack.com/technology/enriching-the-connections-of-the-web-making-the-web-smarter.

Waibel, Günter, and Ricky Erway. 2009. "Think Globally, Act Locally: Library, Archive, and Museum Collaboration." *Museum Management and Curatorship* 24 (4): 323–335. doi: 10.1080/09647770903314688.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

Kathi Martin, Spencer Lamm, Holly Tomren,
Daniel Caulfield-Sriklad, and Nick Jushchyshyn

18 Experiments in High Resolution Imaging for Exhibition and Publication of Historic Fashion: The Drexel Digital Museum Project

Introduction

The Drexel Digital Museum (DDM) project began as an image archive of selections from the Drexel University's Fox Historic Costume Collection and the private collections of noted style icon Iris Barrel Apfel (Sylvia 2016) and Philadelphia designer fashion retailer Toby Lerner (Wellington 2007). It has evolved into a platform for experimenting with new media production for exhibition and publication of historic fashion. Pursuit of best practices for digital representation and description of objects, and sharing of data helps insure that in the future the researcher will be able to bring together data from a variety of historical sources. Working with the Drexel University Libraries (DUL) team we have developed a strategy to provide description and discovery of the DDM's research objects across multiple platforms, from digimuse, the online presence for DDM to Drexel's iDEA repository.

The complex problem of preserving and disseminating cultural heritage requires a multidisciplinary collaboration of museologists such as (1) historians and makers of cultural artifacts, (2) informatics and metadata specialists, and (3) digital media experts. The Drexel Digital Museum (DDM) project is an international, interdisciplinary group of researchers focused on the creation of new media for exhibition of cultural heritage objects. Collaborators include (1) Drexel University's Colleges of Media Arts & Design, (2) Computing and Informatics, (3) Arts and Sciences, (4) the Drexel University Libraries' (DUL), (5) Islandora based repository, (6) iDEA, (7) the Fox Historic Costume Collection (FHCC), (8) Seoul National University, (9) the Ecole polytechnique fédérale de Lausanne, (10) the Fulbright Foundation, and (11) the Costume Society of America.

Articles of dress become cultural artifacts whose status is conferred by a social configuration of aesthetics, consumption, class, capital and personal identity. Historic changes in the collection, exhibition, and dissemination of fashion present multifaceted perspectives on how exhibition and publication of fashion

https://doi.org/10.1515/9783110337419-019

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

can be used to reveal the ideas and concepts of societal processes (Melchior 2014). The DDM is researching the production, curation and publication of high resolution, interactive, panoramic images of fashion objects to integrate into new modes of display and publication. The stories we create from these assets are freely disseminated via the Internet and a trusted digital repository. Best practices for producing innovative media to exhibit and publish historic fashion are being employed throughout the project.

The Drexel Digital Museum Project

Visiting crowds at the highly influential 1876 Centennial Exposition in Philadelphia, were edified by museums bringing together representations of cultural heritage from around the world and making them popular entertainment, designed to elevate society through the morally beneficial influence of great art (Conforti 1997). Driven by a complex set of motivations, ranging from civic duty and personal recognition, to the transfer of the homogeneous values of America's elite to the country's broad, mostly immigrant underclass, A. J. Drexel founded The Drexel Institute of Art, Science, and Industry (Rottenberg 2006). Understanding the influence and inspiration of both fine and design arts along with the study of science and industry in preparing young men and women for the work force, Drexel instructed the Institute's first president, James MacAlister, to travel to Europe, and later, to the Chicago World Fair, to purchase design and fine art from which the Drexel Collection was born. When The Drexel Institute opened its doors in 1892 these purchases, including textiles and non-western dress, were on display in vitrines on the first floor of the Institute to inspire the students and the visiting public. The students collected engravings and plaster casts, reproductions of works of art in other collections, for study. In the 1920s, specimens of European and American dress were added. The Drexel Historic Costume Collection, now the Robert and Penny Fox Historic Costume Collection (FHCC), has continued to grow to a current estimated 14,000 fashion objects and is noted as one of the finest fashion-teaching collections in the United States.

In the mid-1990s historic costume/fashion collections such as (1) The Costume Institute of the Metropolitan Museum of Art, (2) The Kent State University Museum, (3) The Fashion Institute of New York, (4) The Texas Fashion Collection at the University of North Texas, and (5) the FHCC began updating and standardizing their archives. The Drexel Digital Museum (DDM) project began as an image archive of selections from the FHCC and the private collections of

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

noted style icon Iris Barrel Apfel (Sylvia 2016) and Philadelphia designer fashion retailor Toby Lerner (Wellington 2007). The mission of the DDM is to provide open access to fashion design collections and exhibitions for scholarly research and enjoyment by the general public. Through the Internet component, DigiMuse, users can search the archive and retrieve high quality, panoramic images of fashion objects from the database. Or, they may roam galleries of fashion such as those held in the Pearlstein Gallery of the Antoinette Westphal College at Drexel University.

Reproduction of art and design in the digital age

Photographic surrogates were quick to replace the engravings and plaster casts that Western museums used to represent, publicize and share their holdings for study. Inexpensive digital imaging and the Internet have radically changed how images and information about cultural artifacts are disseminated. In the mid-nineties, online resources were charged to enhance learning by providing travel through simulated galleries where the viewer could manipulate the objects and have immediate access to catalog data and information to contextualize the object (MacDonald and Alsford 1991). Collections accessed on the Internet can now transcend the limits of place and time, allowing visitors to meander through displays and interact with objects in previously unimagined ways. As new technologies in imaging emerge, a challenge to humanities collections looking to digital curation and dissemination of their holdings now is to make the quality and significance of the virtual representation meet or exceed that of the original object and to create user interactions which are "conducive to information foraging" by integrating diverse interactive media in presentation of subject matter (Milekic 2007, 382). Most museum directors and curators have adopted new interactive technologies to contextualize and enhance their exhibits, and to widely disseminate images and information about their holdings.

The founders of the high resolution imaging company Factum Arte, Sattin and Lowe (2015), argue that original objects have, through reproduction, many instances of life and "facsimiles, and especially those relying on complex (digital) techniques, are the most fruitful way to explore the original and even to help re-define what originality actually is" (par. 8). Adhering to best practices in reproduction increases the value of the original object by producing multiple quality copies, which can be distributed via the Internet to a much larger audience than can attend a museum exhibition. This democratizes the museum experience and can elicit varied interpretations of the fashion object's status. Our

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

responsibility in the image capture process is to confirm the quality and authenticity of the data used to create the virtual representation. Through selection of the object, lighting to enhance detail, maximizing resolution, and implementing a semantically based knowledge system, such as Metadata Object Description Schema (MODS) mapping, to help to insure perpetual digital discovery and access we become mediators between the object and audience.

A. J. Drexel's vision for a learning environment that would be a confluence of art, science and industry continues as a guiding force in the Drexel Digital Museum project. Design of the DDM began with a survey and interviews of fashion scholars and curators (members of the Costume Society of America) and fashion designers, faculty and students. The results revealed that what the user wanted most from an online resource was (1) high quality images of the garment from multiple views, and of details of construction and embellishment, and (2) multiple search criteria. Initially, the DDM adopted QuickTime Virtual Reality (QTVR) software to provide multiple views and hotspots of details of historic fashion objects. In its initial iteration, the DDM adopted (1) best practices at the time for naming convention from the Computer Interchange of Museum Information (CIMI) (Bearman and Perkins 1993), (2) standards for image quality and processing from the Museums and the Online Archive of California (MOAC) (Rinehart 2003) and, (3) data standards for interoperability through metadata exchange from the Open Archive Initiative Protocols for Metadata Harvesting (OAI-PMH) (Arms et al. 2003).

The DDM and its online presence, digimuse, have evolved into a platform for experimenting with new media production for exhibition and publication of historic fashion. As bandwidth speeds have increased, QuickTime Virtual Reality (QTVR) has been replaced with video. Furthermore, the DDM has been researching best practices for utilizing three-dimensional (3D) Object Virtual Reality (ObjectVR) and HTML5 to publish interactive panoramas of historic fashion and to incorporate evolving metadata standards to manage collections of these artifacts. Work to date from DDM includes the creation of prototype 3D interactive high-resolution images of historic fashion and spaces and rich metadata descriptions of these artifacts. The resulting files will be used in exhibitions, displayed online, and accessed and published from Drexel University Libraries Islandora-based repository, iDEA. Drexel's iDEA repository provides permanent, open access to the unique digital resources produced by the Drexel community.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

Image capture

To produce a 3D image of a fashion garment we must first reproduce the three-dimensional object through a series of two-dimensional (2D) digital images. It is critical to preservation that the 2D images meet the highest standards for archiving and accessibility. Technical standards for quality image capture are constantly evolving and no universal standards currently exist. We looked to the best practices recognized by leading governmental institutions such as the *Guidelines for the Creation of Digital Collections* (Consortium of Academic and Research Libraries in Illinois 2014) and the *Technical Guidelines for Digitizing Archival Materials for Electronic Access* (Puglia, Reed and Rhodes 2004) and found they were already in use by the DDM. The standards used in the beginning of the project such as the Open Archive Initiative Protocol for Metadata Harvesting (OAI-PMH) (Martin 2001) have demonstrated their sustainability.

Sattin and Lowe (2015) recommend creating images at optimum resolution and in the raw formats used at the DDM. This ensures our data will be able to be repurposed by future advances in technology (Sattin and Lowe 2015). At the present time, the DDM uses GigaPan technology in order to create high resolution images of 3D objects for the DDM virtual reality iDEA website. Then the GigaPan photographs are repurposed with HTML5 which is a markup language for structuring and presenting multi-media and graphic elements online (Rowinski 2014). The final image capture, now in the form of a three-dimensional object capable of being housed within the iDEA repository is capable of being displayed at up to three times life size, rotated in 360 degree panorama, zoomed into rich detail, and is compatible for online consumption via HTML5.

Image capture process

Under supervision from the curator, the garment is mounted on a mannequin of appropriate size. The mannequin is then centered on a motorized Kaidan rig which will turn the object as it is being photographed and therefore suitable for taking photographs from multiple views. These multiple views allow for the creation of three-dimensional objects. At the DDM, objects are photographed from eighteen different views. It is important that the appropriate lighting is set for all eighteen views. This is because when working with the garments, which are often translucent and/or shear textiles, they are best illuminated with soft light from many directions through the use of large soft boxes and shoot-through umbrellas to illuminate the background. Four strobe lights are synched to the

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

shutter to prevent harsh shadows that would result from single-strobe, focused light. Two strobe lights are used for the background and two for the subject and are adjusted to the height of the camera and the garment. This is important because if the lighting is too high with respect to the subject and camera, downward shadowing and excess contrast results. As such, adjusting the height of the strobes to coincide with the bust of the subject and height of the camera results in a more uniformly lit image. The camera, mounted on a GigaPan robotic head, is positioned sixteen feet from object and four-feet-seven-inches above the ground with the camera lens set to 200mm (locked). The rig is set to rotate twenty degrees with each image capture, giving us eighteen views of the garment. The GigaPan robotic head is set to capture three columns of twelve rows of images for each view, producing 36 Adobe Camera Raw images per view for a total of 648 images per garment.

The raw image files are batch converted to Tagged Image File Format (TIFF) files using Adobe Camera Raw 9.1 in Adobe Photoshop Creative Cloud 2015. PTGui Pro, an advanced panorama photo-stitching program for Windows and Max OS-X is used for the stitching process. PTGui Pro allows for full manual control over the stitching process, which is necessary when using an automatic stitching action. When the TIFF files are loaded into PTGui Pro, the software aligns the images automatically using the Exchangeable Image File (EXIF) data from the camera. The alignment process generates control points, which are matching points on two overlapping images. Most of the automatic stitching process results in a seamless stitch; however on occasion it is necessary to manually create control points when there is no pattern recognition between images. Once the software has aligned all images and enough control points have been determined, the panorama is created as a single image. The exported image is set at the optimum pixel size in TIFF file format. This process is repeated for each of the eighteen views of the garment.

Next, each stitched image is imported into Adobe Photoshop where the garment is separated from the background. The background is then deleted and replaced with a gradient, mid-tone grey backdrop created in Photoshop. This background was tested on a variety of fashion designs to determine a uniform background, which would best display the variety of colors and textures of fashion garments. Object2VR by Garden Gnome Software generates object Virtual Realities (VRs) as QuickTime VR (QTVR). Adobe Flash and HTML5 output formats are then used to turn the TIFF files into interactive 360 degree object movies. The resulting high resolution movies can be displayed up to three times life size and zoomed in for details of embellishment and construction.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

Data standards

The pursuit of best practices for description of objects and sharing of data helps ensure that in the future the researcher will be able to bring together data from a variety of historical sources. Our research on classification structures revealed the development of new standards and strategies for sharing information between database interactive websites. Early in its digital strategy, the Library of Congress stated, "the Library should selectively adopt the portal model for targeted program areas. By creating links from the Library's website, this approach should make available the ever-increasing body of research materials distributed across the Internet" (National Research Council 2000, 92). We chose to incorporate the standards of the Open Archive Initiative Protocol for Metadata Harvesting (OAI-PMH) to enable this distribution.

eXtensible Markup Language (XML) provides rules by which communities from various domains may transport data over the World Wide Web. This extensibility allows communities to create their own sets of tags. The logical grammar of Document Type Definition (DTD) defines these tag standards to the greater community. Using agreed upon protocols for the tags allows for the harvesting of metadata by a variety of browsers. Resource Description Framework (RDF) can be used to enhance Dublin Core (DC) defined markup by supplying a framework for expressing relationships among items, sets of items or entire collections (Martin 2002).

In the two-party model as described by Lagoze and Va de Sompel (2003), data providers and service providers use Hypertext Transfer Protocol (HTTP) encoding and XML schema for protocol conformance. Extensibility is achieved by providing item and collection level metadata. This allows for searching for an item by descriptive fields like object title, object creator, etc., or by the collection to which the object belongs. The tags for the Open Archive Initiative Protocol for Metadata Harvesting (OAI-PMH) are divided into three sections: protocol support; format-specific metadata; and community-specific record data. The sets or collection definitions are defined by the communities of the data providers and are not defined by the OAI-PMH protocol. The <title> namespace in the format-specific metadata allows different domains to have different meanings for the XML components (Lagoze and Van de Sompel 2003).

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

Metadata – Original object

Since we did not inherit a legacy system, we decided to create a hybridized cataloging form by adapting existing classification structures for art images, fashion, and textiles, to the needs of our users. The Art Museum Image Consortium (AMICO), The Research Libraries Group (RLG), the Dublin Core Metadata Initiative (DCMI), and Collections Trust (previously the Museum Documentation Association) in the United Kingdom have been actively involved in guiding metadata communities to document, describe, and manage museum objects and other cultural heritage resources across different domains. For descriptions of the original object, we borrowed extensively from the Core Categories for Visual Resources (VRA Core) and the fields used in the Museum Educational Site License Project (MESL), the Dublin Core standard and AMICO's data dictionary allow the project team to specify data elements related to materials, dimensions, provenance and production techniques along with other technical, administrative, descriptive, and preservation metadata elements (Goodrum and Martin 1999).

Controlled vocabulary

In order to organize data about an entity into information readily accessible to all, there first must be an agreed upon name for the entity or a process that can structure a uniform, unique identity for it.

Our user groups included historic costume collection scholars, fashion design students and faculty, and fashion designers, who, we discovered, shared little agreement on fashion category terminology. To incorporate as many varieties of terms, we decided to create a hierarchy of fashion terms that would marry the historic costume collection terminology research conducted by the International Committee for Museums and Collections of Costume of the International Council of Museums with historic and contemporary fashion design terminology from Pickens (Pickens 1957), and the vocabularies of the Getty Research Institute. Synonyms are used extensively as a means to assign multiple terms to the object's identity. In our current hierarchy we require citations from three publications in which the term appears before it can be added to the hierarchy. The publications may be from a variety of disciplines including fashion, art history, cultural heritage, popular culture, etc.

Technology can aid content managers and experts in directing the individualized discovery process through customized descriptive terms and repurposed collections' content (MacArthur 2011). However, opinions vary as to which

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

knowledge authorities and learners are in the museum community. We are now looking to extend this structured vocabulary with a tool that will allow all users to submit terms. Falk and Dierking (2013) suggest a scenario "through which the strengths of both the institution and the visitor are the basis for new knowledge ... co-created knowledge and meaning making" (299). In the next iteration of our taxonomy all users will be able to submit new terms and synonyms for existing terms. As the synonyms for one term are used in a search, the term will ascend in the order of the synonyms. In this structure, we maintain the authority of the experts at the top end of the hierarchy and extend authority, through crowdsourcing, to the user population. Through the inclusive human connection in this exchange of authority we will use crowdsourcing to develop trust in the classification system. We hope to attract the interpretive community, fashion specialists, and the general public who will bring their own historical, aesthetic and cultural references to the vocabulary.

Metadata – Digital object

The Drexel University Libraries (DUL) team has developed a strategy to provide description and discovery of the DDM's research objects across multiple platforms, including iDEA. While the Object Virtual Realities (ObjectVRs) will be hosted and accessible via digimuse, the record of the efforts to create them will be preserved and made accessible via iDEA. The ObjectVRs are "stitched" together from hundreds of still images. Those still images, as well as documentation of the process used to stitch them into ObjectVRs, will be recorded in iDEA. A key focus of iDEA is to serve as a repository for the research output of the institution, both for preservation and to provide the opportunity for other researchers to examine and repurpose the process and results of research. By providing a permanent record of the DDM's workflows and content, iDEA will allow others to go beyond interacting with the ObjectVRs and allowing them to understand how the objects were created.

To provide a framework for this, DUL has worked with the DDM team to map metadata from the customized fields in digimuse to Metadata Object Description Schema (MODS). MODS is a robust metadata schema maintained by the Library of Congress which provides a high degree of granularity and extensibility and can be used to describe a wide variety of physical and digital objects (Guenther 2003). MODS also supports the use of embedded uniform resource identifiers (URIs) for entities such as names and topics, which will facilitate eventual exposure of these assets via linked open data. Islandora's open source

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

software framework supports multiple metadata schemas, but DUL has worked to position MODS as the central schema in the system, and based other schemas and application functions on the MODS implementation wherever possible.

The teams worked collaboratively to map each individual data element, a process that produced engaging discussions and compromises. For example, we grappled with questions about the granularity of description, since the compound objects in iDEA require describing the images as a set, rather than describing each individual file. "MODS should complement other metadata formats and should provide an alternative between a very simple metadata format with a minimum of fields and no or little sub-structure (for example, Dublin Core) and a very detailed format with many data elements having various structural complexities such as MARC 21" (Guenther 2003, 139). In addition, genre and format-related fields were particularly challenging to map, as MODS provides a variety of options. However, after several iterations of the mapping, the teams reached a compromise that satisfied our needs.

In order to permanently connect a digimuse ObjectVR with the related objects in iDEA the DUL and DDM teams have also agreed to use the Persistent Uniform Resource Locator (PURL) generated by iDEA (Handles) as well as the Libraries' subscription to the EZID service from the California Digital Library. iDEA's handles will be used to provide a permanent link from the ObjectVR back to the related content in iDEA. Similarly, EZIDs will be applied to all ObjectVRs and will be added to the metadata for objects in iDEA as a way to link from the research to the final product. In this way, the research and its output will be tied through PURLs and will provide a reliable mechanism for directing users from one to the other regardless of their starting place. The PURLs will also be useful for our future goal to eventually release this collection as linked open data, a best practice for exposing, sharing, and connecting pieces of data, information, and knowledge on the Semantic Web. In order to facilitate the sharing and reuse of the content in digimuse and iDEA, exploration has begun into the use of a Creative Commons license.

The digital exhibition

Museums have been experimenting with interactive digital media, merged with display of the original object. Interactivity is used to create a participatory visitor experience, which responds to users' various learning styles, interests and knowledge (Fry and Holland 2013). If the quality and interactivity of the digital replica are of the highest caliber and captivation, could the replica take the

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

place of the original for exhibition? Particularly when the original is fragile and finite?

It is very expensive to dress, ship, insure and prepare displays of original historic fashion objects. The objects must rest for at least three years between exhibitions and will deteriorate with each additional exhibition. Our high quality, interactive ObjectVRs record the original object at that moment in time, before further deterioration. The higher the amount of data (resolution) we can record, the closer the reproduction comes to the original. High resolution copies of an original can allow the viewer to witness "the many layers of its past including its conservation history, its present state and even perhaps anticipating how the object will be interpreted in the future" (Sattin and Lowe 2015, 1). The digital museum can be loaded onto a computer drive and shipped around the world, scaled up for exhibition on twelve foot monitors or scaled down for 24-inch monitors. The team plans to integrate the ObjectVRs into 3D panoramas of historic spaces and exhibit them in high resolution, large-scale display in a future exhibition in Drexel's Pearlstein Gallery. Using an iPad, the audience will be able to rotate the garment to all sides, zoom in on details and access linked data from other repositories.

Cultural heritage artifacts carry historic knowledge. In *Museums in the Digital Age*, Susan Bautista (2014) acknowledges George Hein's (1995) proposed constructivist museum wherein the museum visitor constructs meaning based on their knowledge and experience rather than passively accept what is imposed by curators and educators. She posits that technology can facilitate this participatory culture (Bautista 2014). The exhibition visitors' ability to interact directly with the digital heritage artifact takes them from a reactive museum experience to active participation in the distributed exhibition, customizing and re-contextualizing their encounter with each object in their experience of fashion. For the most egalitarian dissemination, the digital artifacts are being repurposed for the Internet, where they can be freely and inexpensively shared, numerous times, to a varied audience, singularly or in concert with friends and/or strangers, without harm to the original.

Conclusion

The DDM will provide, via Drexel's digimuse and iDEA, broad access to fashion's rich cultural heritage. The best practices informatics (controlled vocabulary, metadata, PURLS) ensure open and sustained access to the hundreds of images required to generate ObjectVRs as well as the processes that created

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

them. This will allow historians, scholars of material culture and fashion design to access the full-scale 3D images from the iDEA repository with the potential to be recombined for digital exhibition with similar assets from other collections. By emancipating historic fashion from the restrictions of conservation and curatorial practice we hope to transcend place and time by engaging the participant in an interactive, individualized, and visually lush experience of dress. In MacArthur's words, "Let us view the digital revolution not as the death knell for the museum as we know it but as an opportunity to enhance the relevance of our collections in the lives of the public that we serve" (2011, 65).

References

Bautista, Susan Smith. 2014. *Museums in the Digital Age*. Lanham, Maryland: AltiMira Press.

Bearman, David and John Perkins. 1993. "Standards Framework for the Computer Interchange of Museum Information." *Museum Computer Network*. http://cool.conservation-us.org/byorg/cimi/cimifram.html.

Consortium of Academic and Research Libraries in Illinois. 2014. "Guidelines for the Creation of Digital Collections Version 2.0." cdlib.org/services/access_publishing/dsc/contribute/docs/cdl_gdi_v2.pdf.

Conforti, Michael. 1997. "The Idealist Enterprise and the Applied Arts." In *A Grand Design: The Art of the Victoria and Albert Museum*, edited by Malcolm Baker and Brenda Richardson, 23–47. New York: Harry N. Abrams.

Falk, John H. and Lynn D. Dierking. 2013. *The Museum Experience Revisited*. Walnut Creek: Left Coast Press.

Fry, Emily B. and Jennifer Holland. 2013. "Remix: Design, Media, and Shaping Experiences." *Exhibitionist* (Fall): 54–58. https://www.name-aam.org/s/12-EXH-f13-Remix_Holland.pdf.

Goodrum, Abby, A. and Kathi Martin. 1999. "Bringing Fashion Out of the Closet: Classification Structure for the Drexel Historic Costume Collection." *Bulletin of the American Society for Information Science and Technology* 25 (6): 21–23. doi:10.1002/bult.133.

Guenther, Rebecca S. 2003. "MODS: The Metadata Object Description Schema." *Portal: Libraries and the Academy* 3 (1): 137–150.

Hein, George E. 1995. "The Constructivist Museum." *Journal for Education in Museums* 16: 21–23. gem.org.uk/pubs/news/hein1995.php.

Lagoze, Carl, and Herbert Van de Sompel. 2003. "The Making of the Open Archives Initiative Protocol for Metadata Harvesting." *Library Hi Tech* 21 (2): 118–18. doi:10.1108/07378830310479776.

Latour, Bruno and Adam Lowe. 2011. "The Migration of the Aura or How to Explore the Original through its Facsimiles." In *Switching Codes: Thinking through Digital Technology in the Humanities and the Arts*, edited by Thomas Bartscherer and Roderick Coover, 275–298. Chicago: The University of Chicago Press.

MacArthur, Matthew. 2011. "Get Real! The Role of Objects in the Digital Age." In *Letting Go?: Sharing Historical Authority in a User-Generated World*, edited by Bill Adair, Benjamin Filene, and Laura Koloski: 56–77. Philadelphia: The Pew Center for Arts & Heritage.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

MacDonald, George F, and Stephen Alsford. 1991. "The Museum as Information Utility." *Museum Management and Curatorship* 10 (3): 305–11. doi:10.1016/0964–7775 (91)90064-X.

Martin, Kathi. 2001. "Data and Image Standards of the Open Archive Initiative: A How and Why for Small Collections." *Conference Proceedings, Cultural Track, W3C 2001*, Hong Kong, China. http://wwwconference.org/proceedings/www2002/alternate/722/index.html.

Melchior, Marie Riegels. 2014. "Understanding Fashion and Dress Museology." In *Fashion and Museums: Theory and Practice*, edited by Marie Riegels Melchior and Birgitta Svensson: 1–16. London: Bloomsbury Publishing.

Microsoft Developer Network. "Get Ready for Plug-In Free Browsing." Revised August 23, 2017. https://msdn.microsoft.com/en-us/en%E2%80%93us/library/hh968248(v=vs.85).aspx.

Milekic, Slavko. 2007. "Toward Tangible Virtualities: Tangialities." In *Theorizing Digital Cultural Heritage: A Critical Discourse*, edited by Fiona Cameron and Sarah Kenderdine: 369–382. Cambridge: MIT.

National Research Council. 2000. *LC21: Digital Strategy for the Library of Congress*. Washington, DC: The National Academies Press. doi:10.17226/9940.

Pickens, Mary Brooks. 1957. *The Fashion Dictionary: Fabric, Sewing and Dress as Expressed in the Language of Fashion*. New York: Funk & Wagnalls.

Puglia, Steven, Jeffery Reed, and Erin Rhodes. 2004. "Technical Guidelines for Digitizing Archival Materials for Electronic Access: Creation of Production Master Files – Raster Images." *U.S. National Archives and Records Administration (NARA)*, https://www.archives.gov/files/preservation/technical/guidelines.pdf.

Rinehart, Richard. 2003. "MOAC – A Report on Integrating Museum and Archive Access in the Online Archive of California." *D-Lib Magazine* 9, no. 1 (January), http://www.dlib.org/dlib/january03/rinehart/01rinehart.html.

Rottenberg, Dan. 2006. *The Man Who Made Wall Street: Anthony J. Drexel and the Rise of Modern Finance*. Philadelphia: The University of Pennsylvania Press.

Rowinski, Dan. 2014. "The Call to Kill Adobe's Flash in Favor of HTML5 is Rising." *Applause*, July 14, arc.applause.com/2015/07/14/flash-vulnerability-html5-replacement/.

Sattin, Anthony and Adam Lowe. 2015. Meet the Master of Reproduction. *Christies Magazine*, November 9. http://www.christies.com/features/Master-of-reproduction-Adam-Lowe-and-Factum-Arte-6776–1.aspx.

Still Image Working Group. 2016. "Guidelines: Technical Guidelines for Digitizing Cultural Heritage Materials." *Federal Agencies Digitization Guidelines Initiative: Creation of Raster Image Master Files*. http://www.digitizationguidelines.gov/guidelines/digitize-technical.html.

Sylvia. 2016. "Style Icon Iris Apfel." *40+Plus: Where Women over 40 Ignite their Style*. May 30. http://40plusstyle.com/style-icon-iris-apfel/.

Wellington, Elizabeth. 2007. "Toby Lerner Passes the Baton." *Philly.com*, January 7, articles. philly.com/2007–01–07/news/25221309_1_european-designers-fashion-boutique

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

About the authors

Christine M. Angel, Ph.D. is Assistant Professor in the Division of Library and Information Science at St. John's University. Dr. Angel earned her Ph.D. and a Certificate in Museum Management from the University of South Carolina in 2012. In 2007, she received an M.S. in Library Science from North Carolina Central University; and, an M.S. in Instructional Technology, a Certificate in Distance Education, and a Certificate in Library Media from East Carolina University.

Shelley E. Barba is the Senior Metadata Librarian and ETD Curator at the Texas Tech University Library. She earned a Masters in Library Science from the University of North Texas, and has published on digital collections management, metadata discovery, and superheroes.

Daniel Caulfield-Sriklad is a creative producer, researcher and lecturer with multidisciplinary experience of physical and digital communication design within the cultural and creative industries. He holds an M.A. in Fashion Curation from the University of the Arts London and an undergraduate degree in Communication Studies from Dublin City University. His M.A. research sought to question the static mannequin display of dress within the exhibition space and look towards digital technology as a means to enhance the communication of the fashion object, specifically looking at communicating the sensorial, tactile and embodied nature of dress. His research led to being awarded a Fulbright scholarship to work with The Drexel Digital Museum Project in 2015. Daniel currently teaches Communication Design for Fashion at the University of the Arts London and his work continues to explore the boundaries of physical and digital space while prioritising the human experience that is central to communication design.

Jennie Choi is General Manager of Collection Information in the Digital department at the Metropolitan Museum of Art. She oversees the museum's collections management system, digital asset management system, and image rights and permissions. Jennie has worked at the Museum since 1995 and has managed the collection database since it was first implemented at the Met. She led the museum's documentation standards committee, has overseen numerous data conversions and upgrades, introduced many improved workflows in administrative functions, and was instrumental in getting collection records onto the museum's website.

Tess Colwell is the Digital Projects Archivist at Brooklyn Historical Society where she manages the existing digital assets, facilitates ingest of born-digital collections, and digitizes two-dimensional collection items. Tess earned her B.S. in Visual Communication from Ohio University, an M.A. in Humanities from Hood College, and most recently an M.L.I.S from St. John's University. She is passionate about photography, history, and Brooklyn.

Dr. Michelle Emanuel is the Head of Cataloging and Metadata Services in the University of Mississippi Libraries. She holds a Ph.D. in French literature and an M.L.I.S. from the University of Alabama, and is the author of *From Surrealism to Less-Exquisite Cadavers: Léo Malet and the Evolution of the French Roman Noir (Editions Rodopi).*

Giovanna Fiorino-Iannace was the Senior Manager of the Antonio Ratti Textile Center in the Metropolitan Museum of Art, where she had overseen the daily operation of the Center's study rooms, storerooms and reference library for more than twenty-one years. She was responsible for the care and handling of the museum's historical textile collection, including over 35,000 textiles, carpets and tapestries belonging to twelve curatorial departments, and provides talks about the facility to students, museum groups and

https://doi.org/10.1515/9783110337419-019

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

visitors. Prior to joining the Ratti Textile Center, Giovanna worked in the department of European Sculpture and Decorative Arts as Research Assistant to the textile curator. She was responsible for the inventory of the textile collection and organized the physical transfer of over 17,000 European textiles, carpets and tapestries to the newly constructed Antonio Ratti Textile Center. One of the projects she coordinated as Research Assistant included the imaging and data entry project for the department, involving the addition of cataloging information and digital images into the museum's collections management system, The Museum System (TMS). She completed both her Ph.D. and M.Phil. in Art History at the CUNY/Graduate Center, and an M.S./L.I.S. from St. John's University.

Caroline Fuchs is Associate Professor and Learning Design Librarian at St. John's University. She holds an M.L.S. from St. John's University, as well as an M.A. in English and an M.A. in history from the City University of New York (CUNY). She is a Senior Fellow at the Vincentian Center for Church and Society, a Center for Teaching & Learning (CTL) Teaching and Technology Fellow, and a CTL Technology Associate. Active in the library community, she is Director-at-Large for the Association of College and Research Libraries (ACRL) and serves on the ACRL Excellence in Academic Libraries Awards Committee. She is also an adjunct professor in the St. John's University Division of Library and Information Science, and is a member of St. John's University Division of Library and Information Science Advisory Board.

Ivey Glendon is the Manager for the Metadata Analysis and Design group at the University of Virginia Library in Charlottesville, Virginia. She holds a B.A. in History and Government from the University of Virginia and an M.S. in Library and Information Studies from Florida State University.

Emily Griffin is a native New Yorker. She received her B.F.A. in Writing, Literature, and Publishing from Emerson College and her M.L.S from St. John's University. While at St. John's University, Emily received the H.W. Wilson Foundation Scholarship for Professional Development. She has been a youth services librarian at Brooklyn Public Library since 2015; it is her first position in the field of library sciences. During her first year of public service, Emily has increased attendance at teen programs at her branch, supervised young adults hired to work at her branch through New York State's Summer Youth Employment Program, and carefully maintained her branch's young adult collection. She's very proud to have served on the first BPL Youth council in 2015. The council aided a teen advisory board in the planning of the first teens-only after-hours event at BPL, an art and performance based celebration called Urban Art Jamm. She looks forward to further serving patrons in her native borough while working at Brooklyn Public Library.

K. Meghan Gross is the former Digital Collections Specialist at the University of Nevada, Las Vegas University Libraries' Digital Collections and was responsible for the digitization and metadata creation for the Southern Nevada Jewish Heritage Project (2014–2016). She is currently the Manager of Digital Asset Services at the Baltimore Museum of Art in Baltimore, Maryland. Meghan has an M.A. in Museum Studies from the George Washington University, and a B.A. in Art History and French from McDaniel College. She is pursuing a Post-Master's Certificate from San Jose State University in Library and Information Science focusing on digital services and emerging technologies.

Michael W. Handis is Associate Professor and Collection Management Specialist at the Graduate School and University Center of the City University of New York. He holds an M.L.S., an M.A. in history and is finishing his thesis for a M.A.L.S. in archaeology. His professional positions include Co-Chair of the Popular Culture Special Interest Group at

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

the Metropolitan New York Library Council; an alternate delegate for the union chapter for his campus; and as a trustee at his local historical society, where he also volunteers. His research interests include ancient libraries and archives, having finished an Omeka site, *Ancient and Modern Libraries: A Comparison* based on his research in Greece while on sabbatical.

Susan Ivey is the Digital Initiatives Librarian and Assistant Professor at the University of Mississippi. She holds a Master of Arts in English from North Carolina State University, and a Master of Library and Information Science from the University of North Carolina at Chapel Hill. She acts on various state, regional, and national committees, including the Mississippi Digital Library as an advisory board member, and the Digital Public Library of America as a Community Representative.

Mike Jones is an archivist, researcher and historian. He is currently undertaking a Ph.D. with the University of Melbourne's School of Historical and Philosophical Studies, supported by the McCoy Scholarship. His doctoral research seeks to explore past, present and future approaches to the management of interconnected artefactual collections and archives in a museum context to support curators, researchers, and the general public. While completing his Ph.D. Mike is also a Research Associate at Museums Victoria, a casual Consultant Research Archivist at the University of Melbourne's eScholarship Research Centre (where he has worked since 2008) and a freelance consultant. Mike's experience includes a Masters by Research in Art History (The University of Edinburgh), extensive experience as a professional archivist, and connections with academic and public library professionals. He has held significant positions on numerous digital and physical archival projects, collaborating with academics, government departments, community organizations and the cultural heritage sector. Mike is also a committee member for the Australasian Association for Digital Humanities and the Victorian Branch of the Australian Society of Archivists.

Nick Jushchyshyn is Program Director, Animation and Visual Effects, Antoinette Westphal College of Media Arts & Design, Drexel University. He has two decades of professional experience in visual effects, administration and technical fields to the Digital Media Department at Drexel University. His experience includes work on major feature films such as "The Girl with The Dragon Tattoo," "I Am Number Four," "Let Me In," "The Last Air Bender" and "The Curious Case of Benjamin Button," which won an Oscar for Best Visual Effects as well as the VES award for Best Single Visual Effect for the Year. Outside of feature films, Nick's freelance background also includes projects in television, commercial and advertising productions. He holds an M.F.A. from the Academy of Art University, and a Bachelor's in Science degree in Commerce and Engineering from Drexel University. In addition to working for studios such as Digital Domain and Dive Visual Effects, Nick has taught digital media and animation at the Academy of Art University and the Art Institute of Philadelphia.

Spencer D Lamm is Manager, Discovery Systems at Drexel University Libraries where he manages the systems department. Prior to joining Drexel, he was a Digital Initiatives Librarian at Swarthmore College and the Technology Coordinator for the Tri-College Consortium of Bryn Mawr, Haverford, and Swarthmore Colleges. He received his M.L.I. S. in 2007 and a B.A. in History in 2005, both from the University of Washington. Spencer's work has been focused on digital collections, institutional repositories, and discovery technologies and has included the development of Tripod, the Tri-Colleges' 'bento-box' style open-source discovery layer.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

Cory Lampert is Associate Professor and Head of Digital Collections at the University of Nevada, Las Vegas (UNLV) University Libraries. Her research interests include linked open data in libraries, strategic planning for digital libraries, and mentorship of new librarians. She received her M.L.I.S. from the University of Wisconsin-Milwaukee and her B.A. in Liberal Arts from Sarah Lawrence College.

Emily Lapworth is the Digital Special Collections Librarian at the University of Nevada, Las Vegas (UNLV) University Libraries. She was formerly the project archivist for UNLV Libraries' Southern Nevada Jewish Heritage Project. She is responsible for overseeing digitization workflows for archival materials. Emily received her M.L.I.S. from Simmons School of Library and Information Science, and a BA in International and Global Studies from Brandeis University.

Michelle Levy is a digital archivist for the Missionary Society of St. Paul the Apostle (Paulist Fathers). She earned her M.S. L.I.S. from St. John's University and has a B.A. in History from St. John's University. She is one of the creators of the Academic Service-Learning blog, Hidden Heritage Collections (hiddenheritagecollections.org); and is the co-author of *Exploring Hidden Collections through Academic Service-Learning* published by The Archivists Round Table of Metropolitan New York, Inc. in the Metropolitan Archivist's winter 2015 issue. Michelle has also co-presented the workshop, *Guerilla Digitization: How to Digitize Your Collection with a Low Budget and Not Go off the Deep End*, as part of The DLIS Professional Development Series. Additionally, she is a member of St. John's University Division of Library and Information Science Advisory Board.

Rachel Lipkin is originally from New Jersey, where she received her B.A. in Publishing Studies at Hofstra University. She then proceeded to work for Macmillan Higher Education, and later found inspiration in the library industry. Rachel moved from New Jersey to California while obtaining her Master's in Library and Information Sciences at St. John's University. While she attended school, she pursued her other passion, baking, and opened her own business, Rachel's Bake Lab. Rachel currently works for Epicurean Group as their Pastry Chef at Sacred Heart Preparatory School. Grateful for her experience with St. John's University, she wishes to continue a career in library studies.

Dara Lohnes-Davies is the Manager of Collections at the University of Wyoming Art Museum in Laramie, Wyoming since July of 2017. She completed her MA in Museum Studies at the University of Leicester in Leicester, England. Her BA is from the University of Wyoming with a major in Humanities and Fine Arts and a double minor in Museum Studies and Music. Previously, she was the Photographer at the National Music Museum in Vermillion, South Dakota. Other museums she has held internships and volunteered are the Oxford University Museum of Natural History, the University of Wyoming Art Museum, and the Herrett Center for Arts and Sciences where she has worked on many digitization and collections management projects. Her specific museum passion is working with databases to further preserve and make collections accessible through that work.

Kathi Martin is Associate Professor, Department of Design, Antoinette Westphal College of Media Arts and Design, Drexel University, and directs the Graduate Fashion Program. She holds a B.F.A. from the University of the Arts, Philadelphia, and an M.S.I.S. from Drexel University. Her fashion design work is included in the Permanent Collection of the Philadelphia Museum of Art and in editorials in *British Vogue*, *Women's Wear Daily*, *The New York Sunday Times Magazine*, *Harpers and Queen* and other fashion magazines. With collaboration from faculty in the College of Computing and Informatics, Drexel University, she has directed The Drexel Digital Museum Project since 1998. She has authored over

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

thirty publications on her research, presented to the international communities of digital cultural heritage, fashion design and historic fashion. From 2010–2014, she was sponsored by a Korean Ministry of Culture, Sports, and Tourism Brain Korea 21 grant to work with Dr. Hyeong-Seok Ko, Seoul National University. At SNU she worked with the Body and Motion Group to create virtual characters for historic fashion. Kathi has co-curated eleven exhibitions of historic fashion at the Westphal College and sponsored visiting scholars to the Department of Design from Egypt, China, South Korea and Great Britain.

Samantha Norling is the Archivist at the Indianapolis Museum of Art (IMA). In this role, she manages over 130 years of institutional records and related special collections documenting the arts in Indiana. Extending her passions for sharing digital collections and raising archives awareness outside of the IMA, Samantha serves as a Community Representative for the Digital Public Library of America (DPLA) and is the current Chair of the Society of American Archivists' Committee on Public Awareness. Samantha earned her M.L.S. and M.A. in History from Indiana University.

Christina Orozco is a digital archivist for the Missionary Society of St. Paul the Apostle (Paulist Fathers). She is one of the creators of the Academic Service Learning blog, Hidden Heritage Collections (hiddenheritagecollections.org). Christina is also the co-author of *Exploring Hidden Collections through Academic Service-Learning* published by The Archivists Round Table of Metropolitan New York Inc., the winter 2015 Issue of Metropolitan Archivist. She has also co-presented the workshop, *Guerilla Digitization: How to Digitize Your Collection with a Low Budget and Not Go off the Deep End*, as part of The DLIS Professional Development Series. Christina earned her M.S. L.I.S. from St. John's University as well as a B.A. in Classical Studies from Hunter College. She is also a member of St. John's University Division of Library and Information Science Advisory Board.

Joy M. Perrin is the Department Head of the Digital Resources Unit at the Texas Tech University Library. She earned a Masters in Library Science from the University of North Texas, and has published articles, book chapters, and a book on the topic of digital libraries.

Dr. Arkalgud Ramaprasad is Professor Emeritus of Information and Decision Sciences at the University of Illinois at Chicago. His current research focus is on the application of ontological analysis and design for envisioning and realizing large, complex systems – especially higher education systems and healthcare information systems. He has a Ph.D. from the University of Pittsburgh, Pittsburgh, PA, 1980; M.B.A. from the Indian Institute of Management, Ahmedabad, India, 1972; and B.E. (Electrical), from the University of Mysore, Karnataka, India, 1970.

Saleha Rizvi works as an Information Technology Coordinator at the University of Illinois at Chicago. As an IT coordinator, she coordinates the technical implementation of campus-wide software used for Academic Assessment, Degree Programs, and Course Processing. She is passionate about using technologies to aid decision making in different domains like library science, health information management, and higher education. She holds a Master's Degree in Computer Science from the University of Illinois at Chicago, completed graduate coursework at Loyola University Chicago, and holds a Bachelor's Degree in Information Technology from Jamia Millia University, New Delhi.

Christopher Starcher is the Digital Services Librarian at the Texas Tech University Library. He earned a Masters in Library Science from the University of North Texas.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

Holly Tomren is Manager, Metadata Services at Drexel University Libraries. She manages a program that promotes discovery and sharing of resources by creating, enhancing, and optimizing metadata for use and re-use in a variety of tools, techniques, and services, in support of the research & scholarly needs of the Drexel University community. She is also a part-time lecturer in Cataloging and Classification at the Rutgers School of Communication and Information. Holly was previously Head of Monograph, Electronic Resources and Metadata Cataloging at the University of California, Irvine Libraries, where her responsibilities included facilitating access to UCI's electronic and digital resources. She has presented about eBooks, cataloging workflows, and mobile access to library resources at conferences including Internet Librarian, Electronic Resources and Libraries, and LITA National Forum. Holly received her Master of Library and Information Science (M. L.I.S.) from San Jose State University and was a 2009 American Library Association Emerging Leader.

Jocelyn Olson Triplett was Manager, Digital Content Management and Dissemination for the University of Virginia Libraries. In this role, she oversees the development of technologies supporting the creation, management, long-term preservation, and dissemination of the Library's rare and unique materials digital collections. Jocelyn has interests in emerging technologies for digital libraries, especially in digital asset management workflows and the dissemination of digitized special collections. She has worked in several departments in the Library as Digital Collections Librarian, Digitization Services Project Manager, and Metadata Specialist. Jocelyn holds an M.L.I.S. from Florida State University, an M.A. in Classics from the University of Virginia, and a B.A. in Classical Languages from Carleton College.

Layna White is Head of Collections Information and Access at the San Francisco Museum of Modern Art, where her department works interdepartmentally to develop the museum's digital records towards maximizing access and sharing for staff and public benefit. The department manages the museum's collection and digital asset management systems; photographs artworks, exhibitions, and museum activities; and manages rights around making and using digital content. Prior to SFMOMA, Layna was Collections Information Manager at the UCLA Grunwald Center for the Graphic Arts at the Hammer Museum, where she was involved in integrating collections needs with on-site and online public access. Degrees in art history and library and information science complement interests in museum practices around documentation and understanding pluralistic and changing needs around access to and use of museum collections.

Travis L. Wagner is a Ph.D. student in The University of South Carolina's School of Library and Information Sciences, where he also received his Master's Degree in Library and Information Sciences. Travis is also a Graduate Certificate recipient within USC's Women's and Gender Studies Department, where he currently serves as a lecturer. Travis also works with moving image archives at the intersection of community archiving and the preservation of obsolete magnetic media, focusing on low cost initiatives to make digitally preserved copies of overlooked materials. The majority of research that Travis focuses on centers around the role emerging information organization initiatives play within cultural institutions, specifically focusing on their role in either limiting or expanding discussions of gender within moving image archival materials.

Sonia Yaco is Assistant Professor and Special Collections Librarian at University Library at the University of Illinois at Chicago. She has a courtesy appointment in the Department of History. From 2007 to 2013, she was the Special Collections librarian and university

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

archivist at Old Dominion University in Norfolk, Virginia. Yaco was previously was the president of Anlex Computer Consulting. She is a regular contributor to scholarly archival journals on topics that include building cultural heritage collections that reflect the diversity of society and innovative methods to improve the discoverability of archives using emerging technology. Recent chapters appeared in *Oral History, Museums, and Education* and in Society of American Archivists' reader on diversity and inclusion, *through the archival looking glass*. She holds a Bachelor's degree in Sociology and Master of Arts degree from the School of Library and Information Studies, both from the University of Wisconsin, Madison.

Jane Zhang is Assistant Professor at the Department of Library and Information Science, the Catholic University of America (CUA). She holds a Ph.D. in Library and Information Studies with archival concentration from Simmons College, Boston, and a joint Master of Archival Studies and Library and Information Studies from the University of British Columbia, Canada. Her teaching and research areas cover records and recordkeeping, archival theory and practice, electronic records and digital archives, and theory and application of information organization. Prior to joining the faculty at CUA, Jane worked at the Harvard University Archives and the University of Calgary Archives.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

Index

https://doi.org/10.1515/9783110337419-020

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:10 AM

Layna White

9 Art Information Culture: Contemporary Disruptions in Museums

Introduction

What *is* art information? It's like the art it describes. Art information has drama, surprise, affirmation, power. Art information can be multisensory, multilingual, multi-authored, fugitive, sparse, questionable, outrageous, nonsensical, bland, humorless, humorous, influential, persuasive, and an array of other characteristics. In the art information culture of a museum like the San Francisco Museum of Modern Art (SFMOMA) we expect art to change, expect art to be experienced differently, to be one thing among many things. We expect art information to change, to be experienced differently, to manifest in words, visuals, and sounds. We expect individuals to need and apply art information differently. This chapter considers such expectations through the lens of needing and applying information around the complex artwork *Predictive Engineering*[3] by Julia Scher, and how such artworks can disrupt art information practices in museums.

Art information is a big topic. Its coverage can encompass cultural, social, scientific, economic, political, and other conditions. Here, the discussion is narrowed intentionally to the information museums record, what they gather, and how this information is presented and related to what is happening or has happened *to* a work, including events both familiar and unfamiliar, and what is happening or has happened *because* of a work, such as the thoughts and actions of individuals involved with a work. The author suggests that the art information culture in museums is a spectacle of informing, describing, and documenting artworks based on active questioning when individuals probe, consider, and experience artworks in changing situations. The chapter begins with an exploration of museum practices and spaces for artwork information, descriptions, and documentation, with particular attention given to how a complex contemporary artwork like *Predictive Engineering*[3] evokes interesting theoretical and practical questions for museums. The author takes the view that active questioning by individuals is central to museum practices of informing, describing, and documenting artworks. In keeping with this view, the chapter reports on a simple think-aloud exercise around the types of questions people ask about *Predictive Engineering*[3]. The chapter closes with a discussion of the art in-

https://doi.org/10.1515/9783110337419-010

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

formation culture in museums and its regard for gaps, sources, and time, with consideration given as to how well information spaces, frameworks, and standards serve the needs of this information culture as well as complex artworks.

The piece itself is made for questions

Predictive Engineering^3 allows for an intriguing study of the kinds of questioning contemporary works can introduce, and what that means for museum information practices and information spaces.

When *Predictive Engineering*^{3}, that is, *Predictive Engineering* to the power of three, was exhibited at SFMOMA in 2016, museum visitors may have read this introduction to the work written by Rudolf Frieling, SFMOMA Curator of Media Arts (2006–present):

> "This surveillance-based installation responds specifically to SFMOMA's architecture and visitor behavior in real time. The work was originally commissioned in 1993 for the museum's home at that time, in the War Memorial Veterans Building. In 1998 Scher updated it for the current location on Third Street and for the web. In this new iteration, Scher has reimagined the work again, adapting it to the current technology landscape – for example, incorporating a drone camera – and with a presence on all floors of the museum's expanded space. *Predictive Engineering*^{3} integrates scenes from the work's earlier versions with new footage and the artist's playful spoken notifications."

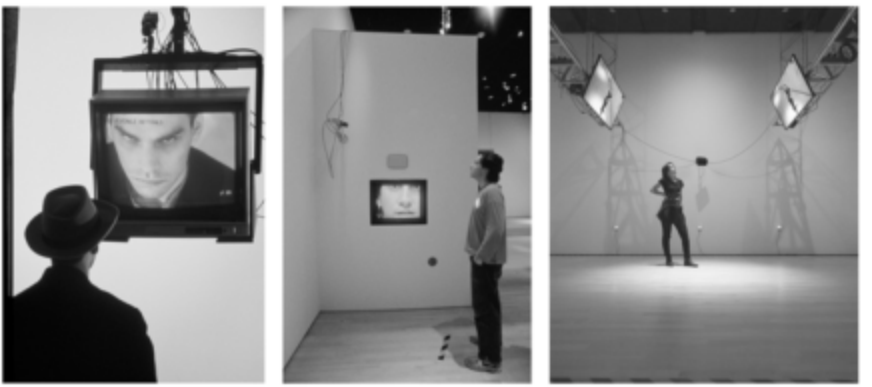

Fig. 9.1: Views of Predictive Engineering to the power of one, two, and three as exhibited at SFMOMA in 1993, 1998, and 2016

Predictive Engineering^3 is situational to San Francisco and its museum of modern art in 2016 (the museum having completed a major building and program

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

expansion in that year). The work is aware of the location, spaces, and people inhabiting those spaces, however briefly. Visitors exploring this surveillance-based installation as it was exhibited at SFMOMA in 2016 would find a mix of contemporary, flat panel display monitors and older, boxier monitors showing changing video scenes of situations taking place in the museum's spaces. These may be scenes with or without people, such as a congested corridor or a vacant window seat, and scenes taking place in public and semi-private spaces within the interior and around the exterior of the museum. Visitors would see video scenes of people behaving in ways that suggest either unawareness or awareness of a surveillance camera, such as scenes with people behaving indiscreetly or simply going about their business, as well as scenes of people performing knowingly to a surveillance camera. This scene-watching would be punctuated by the occasional scene-hearing of vocal tracks of announcements or pronouncements played in the work's gallery. These are curious cautionary announcements, such as "Climax drones are landing on this platform. Attention: climax drones landing on this platform," and what might be received as encouraging pronouncements, such as "This is a space of hesitation. Feel free to hesitate here." The experience is one of watching and being watched. It is a disturbance of scenes, seen and heard, past and present, staged and live. (Appendix B contains a description, or a mise en scène, of what a visitor might have experienced when exploring *Predictive Engineering*[3] as exhibited in 2016.)

Fig. 9.2: Installation view of *Predictive Engineering*[3] as exhibited at SFMOMA in 2016

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

SFMOMA has exhibited *Predictive Engineering*[3] three times to date (1993, 1998, and 2016). Each exhibition has occasioned imagining anew the expression of the work in response to changing situations. For example, the work responds to changes over time to San Francisco and its museum of modern art, to changes in experiences and behaviors with surveillance, and to changes in technology, such as the means of capturing, controlling, and playing video and audio. Scher has described this work as being episodic: first *Predictive Engineering*[3] in 1993, then Predictive Engineering in 1998 and Predictive Engineering in 2016, with Predictive Engineering assumed to follow. The views of the artwork shown in Figures 1–3 suggest the distinct expression of each episode. The work is generative partly in that a new episode can include recordings of scenes from past episodes. That is, Predictive Engineering contains scenes from episodes one and two, along with new recordings of scenes directed by the artist expressly for the new episode, as well as live footage captured during the museum's public hours, as people move through the museum's spaces and experience or play with the work.

Informing

Here, informing is meant as sharing and shaping information, and vice versa. Sharing some of what is known or thought about a work can help shape ideas or decisions, resulting in changed or new information that might in turn be shared. The possibilities for *what is* art information are immense. Asking what art information *is not* will not necessarily help, apart from observing that as a whole it is not static and not indisputable. The most widely shared information about a work is its core information, as typically presented by museums in wall labels situated near artworks on display and in captions accompanying images of artworks in an exhibition catalog or a museum's online collection. The core information for Predictive Engineering3 shared on a label near the work when exhibited at SFMOMA in 2016 read:

> Julia Scher
> American, born 1954
> *Predictive Engineering*[3]
> 1993 – present
> Multi-channel video and sound installation, live cameras, sensors, microphone, mirrors, tape, plastic balls, drone, and text messaging service
> Accessions Committee Fund purchase, 1998

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

Core information has typically been a replicable pattern of artwork information reduced to the most basic of artworld identities. In its most customary state, core takes a few data elements from possible hundreds to help identify works that are created, acquired, exhibited, damaged, or lost. For example, Object ID (2017) is an "international standard for describing cultural objects, in order to facilitate their identification in case of theft" (first par.) and is built largely on core information. The data publishing schema Lightweight Information Describing Objects (LIDO) contains commonly expected data categories that museums may wish to use for artworks, such as Object/Work Type and Title/Name, and presents a view into core information (2010).While the few data elements that comprise core information may be basic (e.g., artwork type and title) determining or agreeing on the data values to be entered into those elements may not be simple because individual or situational preferences bring about a dynamic to expressing art information. The idea of core information can include the desire to have at least one visual or audio representation of the work to aid in identification.

Core information can and does change based on changes in preferences-new understandings, but core is predictably more stable than the wide gamut of information related to what is happening or has happened to, around, or with the work, in events familiar and unfamiliar. Core is the dependable, formulaic music of our art information culture. It is like the recording of a pop song played all summer long. Museums approach recording core information rather clinically and circumspectly, because it is typically the most publicly circulated information about the work. There are norms for recording core information, as seen in museum labels and captions, and perhaps expressed in a museum's internal guidelines or drawn from data content standards such as Cataloging Cultural Objects (CCO) (Baca and Visual Resources Association 2006). Norms which, when observed, can result in core information seeming tightly packaged. Museum norms suggest that core information should be predictably structured and styled so that users can look across information for groups of artworks, or do interesting things with that information, such as showing playful equivalents of physical dimensions for artworks by converting numerical expressions into more relatable descriptive terms (e.g., 84 x 36 inches is roughly the size of a doorway). The "relatable descriptive terms" can be especially helpful. Not the "type of information." which can be especially helpful for people when the work itself is not present.

There is a long-expressed interest in having museums consistently apply agreed-upon guidelines for describing art in order to make it easier to share, find, understand, and use art information. This is perhaps most achievable with core information, but each museum will have its own desire and threshold for

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

consistently applying any universal guidelines. In a somewhat dystopian essay, Pepi (2014) writes of museums adopting the metabolism of a database stating that, "an institution's cultural capital appeals to metrics that emerge from the metabolism of the database: the ability to be queried, manipulated, updated, sorted, and accessed simultaneously" (4).

Encountered in isolation away from a work, the basic artworld identifications understood as core may not be enough for considerable situations and active questioning around artworks, such as when wanting to become more familiar with a work, or when installing a work for the first, second, or third time. *Predictive Engineering*[3] is a striking example of a complex work for which customarily understood core information is immensely insufficient, and for which the shape of core must expand to include information beyond basic identification.

Fig. 9.3: Installation view of *Predictive Engineering*[2] as exhibited at SFMOMA in 1998

Describing

If core information informs circumspectly about a work's basic identity and is packaged for ease of circulation, then making descriptions can be the more

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

elastic expression of ideas, interpretations, or suppositions about the work. Descriptions can be diversified and thereby encompass a fuller embodiment of how individuals talk about, imagine, and remember artworks. The impression of *Predictive Engineering*[3] recorded in the mise en scène in Appendix B is an expression meant to help the reader imagine the artwork. Another example of an expression is the brief description of *Predictive Engineering*[3] written by the artist for placement in a collection catalog published by SFMOMA:

> I wanted to show wires and how things are connected; what happens to your image; who is in control. Like many artists of the time, I wanted to find a way to unravel, to unpack, to deconstruct, to break apart—or to have an audience engage with, even intervene in—the tools and economies of the trade of bodies and images; the commerce of our identities. (Bloch and Stein 2016, 308)

Descriptions are usually made for an audience and for a reason, and are more boundless than the staid core information, or for that matter other kinds of information atomized and entered into distinct data elements. The idea of description includes making and using images and other media to suggest a more dimensional impression of the work, for example to conjure experiencing the work in one's imagination or to stimulate dialogue. In the video *Scher at SFMOMA*, for example, the artist walks viewers through *Predictive Engineering*[3] (Older 2016). Descriptions have the potential to add to or influence how individuals experience the artwork, or what is thought about the work, and perhaps more certainly what is done with the artwork. Descriptions, for example, are more likely to influence inclusion of a work in an exhibition than is an artwork's core information. Describing is the artful, experiential music of our art information culture, similar to a particular recording of a performance by Kendrick Lamar or Ornette Coleman.

Attentions and intentions around describing include making more satisfying descriptions of artworks for persons who are blind or otherwise visually impaired. The Coyote project (2015–present) at the Museum of Contemporary Art Chicago seeks to develop guidelines around describing images, including images of artworks, in ways that intend to be vivid, and satisfying to receive, or, a big improvement over deadpan recitations of basic information about artworks or events. This intention foregrounds the question of what makes for satisfying descriptions of an artwork when the work is absent or not seen, perhaps because an individual is sight impaired, or the work is in storage or installed in a different city, or is lost. Does "satisfying" mean that descriptions will help a person imagine or experience the work in their mind, or help them get a better take on the work, or encourage a person to explore the work, or do something with

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

the work? How to approach making satisfying descriptions of a complex work like *Predictive Engineering*[3]?

Documenting

Documenting artworks is often wrapped up with a changing or changed situation. Some situation or condition *is* happening to or because of the work, or *has* happened to or because of the work, and museums will customarily want to document the situation. It can also be the case that documentation will help museums anticipate what *may* happen to or because of the work. Documenting can be about showing a museum's work with an artwork, akin to showing the moves made to solve a mathematical problem. Museums capture histories, narratives, interviews, correspondence, instructions, reports, and so forth using words, images, models, and sounds. Documentation can capture in-progress actions or perspectives that lead up to a change: examples include a treatment proposal to repair a work, a draft checklist for an exhibition, a forecast of the cost of acquiring and caring for a work, or an assessment of copyrights around a work. Documenting can be about recording a resulting change or action, as in making images of a treated work or images of an exhibition now open to the public, or crafting a purchase agreement to acquire a work, or agreeing to a license to reproduce images of a work. Documenting artworks involves making choices around how to present or represent the artwork or information about the work attuned to the current situation. Documenting is the cover song of the art information culture: similar to tunes like "(I Can't Get No) Satisfaction" or "Cry Me a River" that have been recorded many times, under different situations, interpreted differently by many performers. The author's references to music are intended to be playful. More seriously, the musical score as a documentation model has been considered deeply by Richard Rinehart in his work to describe a notation system for new media art (e.g., see Rinehart 2014). Other references to scores include the artist Rafael Lozano-Hemmer's description of software as the "score" (Lozano-Hemmer, 2015).

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

Fig. 9.4: View near the entry into *Predictive Engineering*[3]*'s* gallery at SFMOMA in 2016

There is usually a situation or condition to be acknowledged when documenting a work, or when using the resulting documentation. For instance, before the museum's photographer makes photographs of *Predictive Engineering*[3] it is useful for the photographer to have at least a general familiarity with the work and understand the purpose of the photographs. Are the photographs intended to document the work's condition during its display, or intended to provide a sense of scale or, as shown in Figure 3, to show visitors exploring the work? When using any resulting photographs or video documentation, an individual's viewing or playback device can influence how one sees the work through the documentation: e.g., monitors or other viewing devices will have different or zero color calibration, and printed materials will reproduce colors differently.

Individuals often rely on others to make or direct documentation. A documenter may bring amazing photography skills and studio equipment to making the documentation or, just the opposite, may bring scant photography skills and limited equipment to the task. The documenter or the requestor may bring a particular editorial style or preference to the documentation. Over the past few years, interest has swelled in museums to have audio-visual documentation that more emphatically shows museums as places where things happen. Museums now make more photographs and videos of artworks with people in and on the scene: for instance, images of people looking closely at works, or videos of an artist working with museum staff to install an artwork in an exhibition. As a

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

surveillance-based installation, *Predictive Engineering*[3] shows the museum as a place where things happen by capturing footage of people exploring or playing with the work. In what ways might SFMOMA direct and stage its documentation of Predictive Engineering? And how might SFMOMA show its work with this work over the past three exhibitions?

Fig. 9.5: View near the entry into *Predictive Engineering*[3]*'s* gallery at SFMOMA in 2016

Surveilling

Museums make locally relevant conceptual dossiers for works. *Dossier* is used here as a nod to *Predictive Engineering*[3] and its concern with surveillance. Anthologies, compilations, or microcosms of information might be other ways of expressing the idea. Museum staff watch and listen to artworks, and to persons making, considering, or experiencing works. Staff dig into logs and data taken from or related to works. Staff analyze, mix, and remix what has been seen, heard, or obtained from the work or from persons involved with the work under changing situations and motivations. This watchfulness is made explicit in, for example, the images and videos museums make of works. The idea of a conceptual dossier intends to be that mix of information, descriptions, and documenta-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

tion about an artwork that is distributed in information spaces throughout a museum. Systems like a museum's collections information, digital asset, and content management systems are art information spaces, as are its art, archive, and library collections, and its analog files related to artworks and exhibitions. The museum too is an information space, as are the individuals who inhabit it, however briefly.

Museums have numerous information spaces intended to help record, gather, and present impressions of active questioning around what is happening and what has happened to an artwork. It is typically a progressive accumulation of contents, and is in some state of being mixed or remixed. Imagined together, these information spaces shape conceptual dossiers with new, old, good, poor, character-rich information, descriptions, and documentation, or impressions of questions posed and to be considered as a result of probing, considering, and experiencing artworks in changing situations. For instance, a note dated 1998 in one SFMOMA analog file for *Predictive Engineering*[3] includes this question posed by staff about the work: "how do we facilitate / administer growth?" This idea of a conceptual dossier is inclusive of the kinds of questions related to artworks that museums like SFMOMA place in mobile apps or in galleries on labels near artworks for visitors to consider. For instance, visitors to SFMOMA galleries in June 2016 may have seen this question printed on a wall near Andy Warhol's *Self-Portrait*, 1967: "What can we decipher from Warhol's Self-Portrait?"

A museum's collections information, digital asset, and content management systems are the most commonly used information spaces by staff. These management systems may have several hundred data elements standing at the ready to receive information about what is happening or has happened to or because of a work. Data elements are standing by to receive information about the artist, exhibitions, publications, conservation, location, ownership, valuation, and so on. Art museum staff might enter information into distinct data elements structured along the lines of a published data structure standard like the Categories for the Description of Works of Art (CDWA). For instance, SFMOMA staff might extract location information from the production schedule used by the artist and the museum when reimagining *Predictive Engineering*[3] for the 2016 exhibition, and enter that information into data elements in a local collections management system: as an example, staff might parse out San Francisco, California into an element akin to CDWA's Creation Place/Original Location.

A well-stocked information management system can help address questions and support thinking about groups of works with similar characteristics, such as, "How many works by women are in this collection?" or "Which are the largest paintings in this exhibition?" Well-used and respected collections information management systems may also help address common questions and sup-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

port interest in events that either have or will happen to individual works such as "Has the museum exhibited this work by Kara Walker?" or "When is the next condition review scheduled for this single-channel video?" Data elements, standing at the ready in these systems tend to be formulaic. Statements are made out of predictable questions: "Who made it?" becomes artist name. "When was it made?" becomes creation date. "Where did it come from?" becomes ownership history. Being formulaic can be useful when working with information across many works, but can be less satisfying at helping or encouraging a person to get to know an individual work. A systematic approach may be fine when the active questioning is fairly simple or core-friendly, but museum management systems can be frustrating to use when questions and curiosities about a work are unpredictable or exploratory, or would benefit from conversational descriptions. What was it like to imagine and install *Predictive Engineering*[3] in 1993, 1998, and 2016? What words are used most frequently in the vocal tracks? What if the individuals who helped build Predictive Engineering are not available to help re-build another Predictive Engineering exhibit?

Ideally, information spaces are designed to be satisfying to use when recording, staging, and reading information about individual works and groups of works. SFMOMA has kept this beacon of fashioning spaces that are satisfying at the forefront of its work with art information. One of SFMOMA's explorations is to test the wiki platform as a potentially interesting workspace for individuals to coauthor, stage, and read information for purposes such as helping the museum keep its work with artworks active, especially in its engagements with artists, research, and public programs. This exploration considers the place of wikis alongside the routine use of management systems for art, digital assets, archives, and library collections. However, gathering contributions of information, descriptions, and documentation from many individuals into current management systems has proven difficult. Wikis are seen today as appealing because the format is familiar and generally likeable, and because the platform has an assumed ease of interaction when coauthoring, editing, and reading. Can the museum fashion wikis, or any other platform or combination of platforms, to be satisfying workspaces to use when questions or curiosities about works are unpredictable or exploratory? Can the museum fashion information practices and spaces to take into account the situational nature of works like *Predictive Engineering*[3], as well as the situational nature of art information? At what point in the museum's work with an artwork, and with an artist, will it show that work by contributing what is known to an information space like a wiki?

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

Predictive Engineering[3] out loud

Individuals go to art information and/or take from art information for a broad range of reasons, and this has significance as to why and how museums record, gather, and present art information. If questions and curiosities about works, and a sense of importance in addressing questions or knowing about answers, are motivations for wanting to share and shape information then, what questions do people have about a work like *Predictive Engineering*[3]?

The author ran a simple exercise around this question about questions. She invited twelve individuals to walk with her through *Predictive Engineering*[3] as exhibited at SFMOMA in June 2016. The twelve invitees each had experience working in museums: some were familiar with Predictive Engineering, others were not. During the fifteen minute walks through the gallery, invitees were asked what they wanted to know about the work. The exercise was intended to be a simple way to gauge the types of questions people have when situated directly in front of exhibited or otherwise installed artworks, and to consider if questions collected from this sample suggest interesting gaps in museum information practices, or suggest anything new. (Appendix A lists the questions posed by invitees.) "What do you want to know?" is a leading question in that it suggests that there are gaps in what a person knows, or that there is curiosity about the work.

Resoundingly, questions offered by invitees in this exercise underscore that people are definitely well beyond being satisfied by core information about artworks. But do the questions suggest anything new about art information in museums? Looking generally at the types of questions asked during the exercise, the author sees questions related to ideas, questions arising from being with the work, and questions for the museum as a steward of artworks and art information.

What is the work? Several invitees asked this question in different ways, probing at ideas related to the work. What about this artwork, and works like it, prompts such a question, and how should our information, descriptions, and documentation help a person explore what the work is? Taking the question literally, our information spaces might contain some combination of the following: impressions describing the work and each of its episodes; lists identifying elements used in the 1993, 1998, and 2016 installations, such as monitors and speakers, video and audio, and elements to be cared for by the museum when the work is not on display; and images or diagrams showing the placement of elements and how people might move through the work's gallery spaces. Beyond a literal reading of "what," works like *Predictive Engineering*[3] suggest we

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

need information practices and spaces that can help convey vividly and satisfyingly "what the work is" by allowing us to synthesize, contextualize, and present information for changing situations.

Fig. 9.6: Watching *Predictive Engineering*[3] video scenes at SFMOMA in June 2016

The *Predictive Engineering*[3] videos for example, received much attention from the twelve invited individuals. *Where are the cameras?* If taken literally, the museum can easily document camera locations with tools such as images and diagrams, and it is assumed that cameras used in the installation can record location data. But invitees were probing the location choices and the types of scenes the locations might yield. Questions such as this might become more challenging to explore the further removed museum staff become from the 1993, 1998, and 2016 episodes, when active institutional knowledge of individuals and

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

spaces has waned. Asking "where" also edges into an interest in understanding how Predictive Engineering works. *What is the artist doing with the recordings?* The artwork has the potential to make a massive amount of video content, between the footage recorded for the staged scenes directed by the artist and the live footage captured during some of the museum's public hours. Asked by invitees with museum backgrounds, questions like this touch on the artist's ideas, as well as what it might mean for the museum's practices in caring for the work. Who is responsible for remembering that footage captured in 2016 should be playable in three, ten, or fifteen years? Can the museum make the footage available to researchers when the work is not installed? How can the museum best describe or document all of the resulting video content, and to what level of granularity? Can the museum learn approaches for description from other industries that make or collect video and audio? The introductory text printed on the gallery wall for *Predictive Engineering*[3], for instance, includes credits for Scher's collaborators in making this episode: e.g., persons responsible for the software and graphics, for the vocals, the photography and editing.

To some extent, invitees may have crafted their questions about *Predictive Engineering*[3] in particular ways because it was a think-aloud exercise. The individuals might have asked different questions if someone they did not know ran the exercise, or if responses could be submitted anonymously. Would invitees have asked the same, or any questions, without the prompt of "what do you want to know?" This simple exercise could be made more expansive and rigorous by including people without museum work experience, or by asking the question when the work is not on exhibition, or by drawing from questions related to why individuals go to art information or take from art information.

What *is* the art information culture in a museum today?

Information practices in museums have come from an information culture that has a regard for gaps, sources, and time. Museums look to their information spaces, or collectively, to conceptual dossiers, to become aware of such interests. *Predictive Engineering*[3] disrupts museum information practices and spaces, as evidenced by the litany of questions it can inspire.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

Gaps

The art information culture in a museum like SFMOMA rewards an inquisitive mindset energized by information spaces that are incomplete and generative. The culture expects fuzziness or unknowns around artworks and thrives on gaps in information spaces created by the works, by persons of influence, and by changing situations. For instance, gaps in what is known or thought about works will be exposed when shaping what will become the thesis of an exhibition. The art information culture thrives on finding gaps in information, in bridging or responding to gaps, or knowingly keeping gaps open, because ambiguity or absence of art information allows individuals to imagine, test, or make new assumptions about the work. Museum management systems support this interest in finding, filling, or knowing about gaps.

The complexity of *Predictive Engineering*^{3} draws attention to sizeable gaps in traditional practices around sharing and shaping information. How best to describe and document the experience of moving through the work when it is on exhibition or otherwise installed? Describe for whom, and for what purpose? Might a three dimensional model or a virtual reality or augmented reality manifestation or tour help fill gaps in *Predictive Engineering*^{3}*'s* conceptual dossier? The author's *Predictive Engineering*^{3} mise en scène (Appendix B) is meant as a sketch for someone who may not have the chance to experience the work in person. The sketch may not be helpful to someone working to re-imagine *Predictive Engineering*^{3} for an exhibition in five or ten years, or wishing to understand the meaning of the work.

How does one best describe and document the work as being re-imaginable in its expression as situations change? *Predictive Engineering*^{3} to the power of three was exhibited in 2016, and the museum can anticipate that there will be a Predictive Engineering to the power of four, and so forth. How does one best communicate questions and plans for managing and caring for this complex work over time? Can SFMOMA look to frameworks like the International Committee for Documentation of the International Council of Museums Conceptual Resource Model (CIDOC CRM) to design ways of expressing what the work is? For example, can the museum use the Conceptual Reference Model to describe the artist's ideas, and the form the work has taken with each episode, as well as what has happened to it and because of it? The museum might investigate the functional units and data entities identified in CIDOC Conceptual Reference Model for their usefulness in describing works that can be re-imagined like *Predictive Engineering*^{3}. For example, CIDOC CRM identifies the functional units Changing Thing and Existence Information, among others, and data entities like

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

Conceptual Object (Beethoven's *Ode to Joy* is given as an example); Propositional Object (plot ideas for films like Kurosawa's Seven Samurai is an example); Information Object (Kurosawa's film *Seven Samurai* is an example); and Production (the seventh state of Rembrandt's etching *Woman Sitting Half Dressed Beside a Stove* is an example).

What words are good for describing what was on display in 2016? Did visitors to SFMOMA in 2016 experience an episode, version, instance, iteration, continuation, re-imagination, or a _____ of *Predictive Engineering*[3]? The standards-bearers of descriptions have been expected to be published content and vocabulary standards. For instance, the Art & Architecture Thesaurus (AAT) is a well-known source for established artwork terms that can be entered into distinct data elements within a variety of museum management systems. However, the AAT does not include the following data values: episodes, instances, iterations, continuations, or re-imaginations. It does however include the data value "version." As such, while the AAT is grand in scope, there are gaps in its application for contemporary art. Others have set out to describe taxonomies for contemporary art. For instance, Freeman et al. (2015) describe a taxonomy for data as an art material, and include terms like Real-time Data in their taxonomy, which might be useful in describing the live feed element of *Predictive Engineering*. See also the Media Art Research Thesaurus guided by Oliver Grau and editor-collaborators. The difficulties of categorizing the practices of some contemporary artists is explored by the Art Genome staff at Artsy in *10 Artists Who Defied Classification in 2015*. It may sound bombastic but, proximity to the ideas around Predictive Engineering may place the artist and museum in a good position to choose words that make the most sense to describe the artwork in a particular context. Word choice can be as situational as artworks, and this makes for lively sparks between descriptions with situational character or gusto, and terms normalized for uses like search, retrieval, and collocation.

Should SFMOMA attempt to characterize some core intellectual platform for *Predictive Engineering*[3] by calling out what the artist, museum, or a community sees as the most important things to know or say about *Predictive Engineering*? Should any conceptual dossier for *Predictive Engineering* have chapters for each episode, and is the sequencing of the episodes in any dossier necessary? Continuing the music analogy, can our information spaces support individuals working like DJs to bring together sets of information, descriptions, and documentation based on changing questions and curiosities?

Scher (2016) has described *Predictive Engineering*[3] as being self-documenting. Playing with this idea, when the work is installed and turned on (as in 1993, 1998, 2016), surveillance cameras are trained on specific spaces around the museum to catch scenes that might be played live in public spaces. Scher

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

might (or might not) use these surveillance videos in future *Predictive Engineering*[3] episodes. In 2016, some museum visitors took pictures while interacting with Predictive Engineering. When *Predictive Engineering*[3] was exhibited at SFMOMA in previous years (1993 and 1998), the museum did not allow photography by visitors in its galleries. *Predictive Engineering*[3] then is documenting the year, the place and space, visitor behaviors, and how individuals respond to being watched, as well as museum policies.

Fig. 9.7: John Baldessari, *A Painting That Is Its Own Documentation*, 1966–1968

Can SFMOMA look to its information spaces, such as its management systems or wikis, to help with gaps in what the museum knows about *Predictive Engineering*[3]? Can the questions expressed by *Predictive Engineering*[3] think-aloud invitees (Appendix A) be explored through the museum's information spaces?

Sources

If the art information culture has protagonists, then art informationists number among them. In a museum setting, art informationists are persons who care about how we express, record, stage, read, and apply art information. SFMOMA art informationists have regularly scheduled forums for considering information and practices around works like *Predictive Engineering*[3]. Our Team Media and Team Architecture and Design are two of our most progressive constructs that bring sustained, shared attention to our work with media, electronic, and/or experiential artworks. The two teams are inclusive of many partners with interests and kinship around the museum's collections, exhibitions, publications, conservation, research, and education programs. The make-up of the teams includes representation from areas of curation, conservation, registration, exhibition, legal, information systems, and public-facing technologies. The spirit and

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

tone established for the forums are an openness and support for pluralistic perspectives and different expertise, and encouragement of active questioning and listening. The teams meet in person regularly and make progress together on immediate and longer-term issues or ideas around complex works like *Predictive Engineering*[3].

Art informationists can go wild with fun or anxiety working with the liveliness of information about an individual work, or a group of works. Art informationists bring the personal into working with information, based on what the informationist knows or wants to say about a work in a particular context. The super-styled core information for *Predictive Engineering*[3] aside, we might expect descriptions and documentation of Predictive Engineering to be as lively as the work itself, and as lively or as individual as the art informationists. For example, visitors to the 1998 exhibition of *Predictive Engineering*[2] at SFMOMA may have read this excerpt from the introductory text written by that exhibition's curator (Robert R. Riley, SFMOMA Curator of Media Arts (1988–2000), printed on the wall near the entrance to the work's gallery:

> "Scher's madcap use of real-time and prerecorded video, provocative spoken-word commands, and original graphics turns the instruments of surveillance outward to the museum as a social and authoritative agency, yet ultimately and comically inward on itself. The surveillance industry, Scher suggests, reflects society's disposition while permeating the individual imagination so deeply as to provide an occupation for it. The artist uses the medium to stimulate voyeuristic curiosity as well as a means to antagonize incongruent social conditions she perceives: indifference to search and disclosure tactics and attraction to customized forms of surveillance that elicit a calculated response from its subjects. *Predictive Engineering*[2] is presented at a moment that coincides with a crisis in culture induced by public disclosure of private matters. The artwork critically apprehends the hard targets and soft applications of surveillance that might otherwise, in their ubiquity, remain invisible."

Art informationists are concerned with the vitae of art information. That is, informationists are interested in knowing the source and motivation for the information, for instance, who made it, when, and for what purpose. There is a strong interest in getting information about works from people with proximity to artists, artworks, and ideas around works. The art information culture can be hungry for fresh interpretations and insights into the work, and museums benefit from multiple perspectives contributing information, descriptions, and documentation to our information spaces, as a kind of checks-and-balances to what might otherwise be homogenous information about a work. Museums are more and more tuned into expectations for livelier, more expansive participation with art. Less charted territory in museums is to get information about works from museum visitors, non-museum information users, and programmers. For exam-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

ple, visitors to the 1998 exhibition of *Predictive Engineering*[2] at SFMOMA may have read this excerpt from the introductory text written by that exhibition's curator (Robert R. Riley, SFMOMA Curator of Media Arts (1988–2000), printed on the wall near the entrance to the work's gallery:

> "Scher's madcap use of real-time and prerecorded video, provocative spoken-word commands, and original graphics turns the instruments of surveillance outward to the museum as a social and authoritative agency, yet ultimately and comically inward on itself. The surveillance industry, Scher suggests, reflects society's disposition while permeating the individual imagination so deeply as to provide an occupation for it. The artist uses the medium to stimulate voyeuristic curiosity as well as a means to antagonize incongruent social conditions she perceives: indifference to search and disclosure tactics and attraction to customized forms of surveillance that elicit a calculated response from its subjects. *Predictive Engineering*[2] is presented at a moment that coincides with a crisis in culture induced by public disclosure of private matters. The artwork critically apprehends the hard targets and soft applications of surveillance that might otherwise, in their ubiquity, remain invisible."

Art information can be respected and used, or dismissed, based on the source and motivation. Interpretations perceived or assessed to be less misfires or important at the time may get less airplay now in the museum's information spaces. While the art information culture in museums is tolerant of gaps, it can be less tolerant of perceived mistakes in information of the should-have-known-better variety that are made public by the museum, especially when the mistaken information could have been drawn from a reliable source. As a matter of degree and context, the consequences of should-have-known-better mistakes in information made public by museums can diminish reputations and trust. Certain information rituals are practiced in the run-up to making information public. For example, preparing for the opening of an exhibition will likely occasion the making of new introductions to artworks for use in wall labels. There is a ritual of validating ideas by surrounding oneself with things and people related to the artwork. And when a work is not yet installed or is otherwise unavailable for viewing, exposure to the work or to the artist may be through information, descriptions, and documentation. There is a ritual of reviewing the currency and/or accuracy of the information at hand, by chasing and recording new information from reliable sources, including the work itself.

Artists, their studios, and/or artist representatives are among the key sources of information, and influence how museums working with contemporary art present and represent a work publicly in museum spaces or museum publications. Museums working with contemporary art can be in a position to engage tangibly with artists about their works. This can be in the course of presenting works publicly and caring for works generally. It can be in the course of devel-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

oping new scholarship around art and can be about staying immersed in the art of the times. For contemporary art museums like SFMOMA, long-term engagements or interactions with artists can lead to meaningful intellectual and creative work by art informationists. Julia Scher and staff at SFMOMA have engaged tangibly while researching, making, and exhibiting each *Predictive Engineering*[3] episode. With *Predictive Engineering*[3], and with funding from The Andrew W. Mellon Foundation, Scher had a formal residency at the museum, with on-site engagements over the course of 2014–2016. The residency involved the artist and staff working together in multi-session, intra-disciplinary dialogues and actions around re-imagining the work for the changed conditions of 2016.

Art informationists have a regard for knowing the context or condition within which a source has made the information, or description, or documentation, so as to understand intentions and surrounding influences. Context can matter because an art informationist brings what she knows, her interests, her temperament and preferences, and her audiences into how she approaches informing, describing, and documenting. An unknown context is one of the gaps wanting to be filled by art informationists, sometimes aided by informed suppositions. Workaday examples of gaps in context are finding images of a work and not knowing when or why the images were made and finding descriptions of a work and not knowing the intended audience.

Time

The art information culture in museums can be restless and restful. Restlessness has been the louder, greedier characteristic of the art information culture. Gaps in information get attention; events get documented; ideas get shaped and reshaped. Museums have reason to record and make information available in good time. For example, pressing questions need answers now; future selves will need to recall decisions made today; and, importantly, access to works will be through recorded information when works are not on display. More restfully, a museum's experiences with and understandings of works happens over time, with different surrounding influences. There may be big gaps in the times between experiences, and therefore museums have reason for taking the time needed to shape and present new information or thoughts about works.

Museums contend with the challenge of art informationists postponing to another time making their information, descriptions, and documentation about works available in common information spaces. Competing demands on time will be a factor in delaying contributions, as well as perhaps a self- or peer-im-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

posed expectation that information should have a certain level of finish or polish before it is contributed to information spaces. This may be true even with the shared understanding that information can be changed or updated, and true even when the information can be wrapped with context and its availability is limited to staff. The art information culture can be a culture of deliberation, with aversion to some risk when museums believe the information may become public. It may be that information spaces have been set up to be, or are perceived to be, spaces of restraint and deliberation. Another possible aspect of postponing the sharing of information is that museums are often both actively studying and actively writing about a work. There is a case to be made for having common information spaces in which informationists can feel comfortable sharing active, sketchy, or tentative information about works, to the benefit of active questioning and curiosities.

One of the most active times for the art information culture is when a work will be or is currently on display. From the planning stages of an exhibition through its closure and thereafter, information about the work is needed, generated or updated, shaped, and shared. Exhibiting a complex work like *Predictive Engineering*[3] generates a great deal of fresh information about its installation, how it works and is working, and how people experience the work in both expected and unexpected ways. Does the art information culture's concern for gaps, sources, and time help or challenge a person's ability to become more familiar with *Predictive Engineering*[3] while it is on display in 2016? SFMOMA staff know from experience that another cause of delayed contributions is that staff are asked to add their information to management systems, such as collections information management systems, that they simply do not like to use or deem as to complicated. At SFMOMA, we are actively imagining lively information spaces through which staff and external collaborators can communicate about artworks in ways that might slip easily into the work rhythms of time conscious, situation aware, art savvy informationists, perhaps through some combination of information, asset, and content management systems, wikis, websites, and other tools.

Conclusion

The art information culture in museums is like the art it describes. It is a disturbance of scenes, seen and heard, past and present, staged and live. It is like the individuals who describe the art, in that there is a disturbance of experiences, preferences, tolerances. This chapter narrowed the big topic of art informa-

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

tion to information that is related to what is happening or has happened to, around, or with the work, while looking most closely at the information, descriptions, and documentation that could be a museum's contribution to what is known about a work.

Art information can be multisensory, multilingual, and multi-authored, and museums benefit from multiple perspectives when sharing and shaping information about artworks. At a museum like SFMOMA, our work with art can benefit from lively, expansive participation with information, where the tone established for participation is an openness and support for pluralistic perspectives and different expertise, and encouragement of active questioning and listening. Our work with art can benefit from information spaces that are satisfying to use for recording, staging, and reading information about individual works, and groups of works; spaces that take into account the situational nature of artworks and art information. *Predictive Engineering*[3] is a discourse-inspiring example of a complex work that disrupts our traditional information practices and spaces. This is to the benefit of all works in the museum's care, as well as to individuals, art informationists among them, who go to and take from our information spaces in order to become more familiar with a work, share and shape information about the work, or modify the work alongside changing situations. About the artwork's generative and situational nature, a *Predictive Engineering*[3] think-aloud invitee asked, "Is it ever finished?" By the same token one might ask if the museum's work with art information is ever finished, given the generative and situational nature of art information. It is hardly finished on both counts.

Appendix 9.A.

Predictive Engineering[3] Out Loud

June 2016
Question: What do you want to know about the work?
Following are responses from the twelve invitees, grouped loosely by a similarity.

About ideas:
- What's the underlying idea?
- What is the title about?

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

- How did the artist originate the piece? How decide to do this?
- Why is the artist interested in surveillance?
- What does surveillance say about us?
- What's the difference between hidden cameras and performance?
- Why would I want to watch surveillance video in a museum?
- How does – or how much does – the historic documentation play into the current stuff, and why?
- What is the purpose of the historic documentation?
- What is the relationship between the old and the new?
- How much of SFMOMA is the work, and how much of the work is SFMOMA?
- Is it documenting technology over time?
- Is it about using surveillance to watch women?
- Is the artist talking about movement and the body?
- What's up with the naked people?
- What is the significance of the nudes?
- What is the thought process of representing humans some nude some clothed?
- Does PE3 swallow PE2 swallow PE1?
- Where does the piece start?
- Where does it end?
- Is it ever finished? When is it over?
- What is, and isn't the piece?
- What is the work?
- What is the art in here?
- What is PE3 in this room?

About being with the work:

- What's going on?
- How did that happen?
- How does it work?
- How do the parts work together?
- What relationship does the artist have with the engineer who built it?
- Am I doing something to make it work?
- What do the electronics along the walls do?
- Is it set up to make me think I'm doing something to make it work?
- If it's about surveillance, why is a monitor missing? Does it point to earlier iterations?
- Why are these monitors hanging from scaffolding instead of on walls?

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

- How much of the [server closet, hidden from visitors] is part of the expression of the work?
- Are the exposed cables part of the aesthetic, or a gag?
- What do the striped balls do, or mean? Are they used in the videos?
- Is the caution tape a demarcation?
- What does it mean, 'mystery meat being scanned?'
- Do the graphics overlaying videos reference something specific? Or is it meant to be ambiguous?
- Where are the cameras located?
- Where are the cameras?
- Why are certain cameras in certain places?
- How did the artist select the spaces in the museum for surveillance?
- Are the cameras linked to the museum's surveillance system, or did the artist pick the camera positions?
- Am I expected – or expecting – to see my image? Either now or later?
- What is the source of the video?
- What's live and what's not?
- Is any of it actually live?
- Is there a definite pattern to the video? If random, how?
- How can you tell what is old and what is new? What is filmed and what is not? Is it supposed to be confusing?
- Is that the museum? What year?
- How many people from rounds one and two are in round 3?
- Was anyone asked to be in the video because of their acting abilities?
- Did the artist direct people to roll down the stairs?
- How many people are interacting with the microphone?
- What am I supposed to experience?
- Are people experiencing the work in a similar manner?

What now museum?

- What is the artist doing with the recordings?
- Is the current surveillance footage being kept? Can it be used again?
- What's the status of any additional, never-used footage? Is it part of the piece? Part of the story?
- What are we doing with the backstage stuff?
- Are there any rights issues with using surveillance video of people?
- Are there rights issues with getting images of people?
- Is it important that we document the iterations? What would that look like in a database?

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

Appendix 9.B.

Predictive Engineering[3] mise en scène

The following is an exercise in making a description of Predictive Engineering[3]*, and is intended as a sketch of what you might have experienced when exploring Predictive Engineering*[3] *as exhibited in 2016.*

Place yourself in San Francisco in June 2016. It's the summer of the Cleveland Cavaliers winning the National Basketball Association's championship, the Orlando nightclub shooting, and Brexit. Imagine being inside SFMOMA. People are moving casually around the museum's galleries: you hear footsteps, doors closing, people talking. Art is installed on walls, hung from ceilings, positioned on pedestals, and placed directly on floors.

Predictive Engineering[3] occupies the whole of a goodly-sized gallery, and has a number of elements: live cameras, monitors, trusses, cables, sensors, a microphone, audio tracks, mirrors, hazard tape, plastic balls, and a drone. Its gallery is roughly the size of a long, narrow dance studio, or a waiting room at a station. It's sectioned into three adjoining spaces: a large rectangular room is at the center, and two much smaller semi-enclosed spaces flank either side of the center room. The two smaller spaces are roughly the size of freight elevators, and give entry and exit passage into and out of the large center room. Wide yellow and black hazard tape outlines the entire space: the tape is laid precisely on the floor around the edges of the gallery.

Situate yourself a few feet outside the entry to *Predictive Engineering*[3]'s gallery. Look up, and to your right: you'll find the work's first obvious visual element. Far above eye-level a moderately large contemporary flat panel display monitor is anchored to a husky apparatus, a truss, suspended from the ceiling. The screen is placed at the height and position you'd expect to catch sight of a surveillance camera, and in fact a camera is attached to the truss near the flat panel display monitor. The screen shows video scenes playing one after another at a mellow, rhythmic pace, with no sound. The scenes become familiar to you, or somewhat familiar, the longer you watch. You catch scenes of the space and people immediately behind you, and then scenes of other spaces that you guess are part of the museum, and then scenes showing people going about their museum visit or business. You register that people are being watched and recorded. Periodically, the video scenes get interrupted by a solid graphic or by text appearing on the screen: a bold yellow and black hazard-like graphic appears briefly, then the text "Alert Predictive Engineering[3] is currently repairing Predictive Engineering and Predictive Engineering[3]."

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

Look away, and step to your left. You're in the entry passageway to *Predictive Engineering*[3], inside one of the small, semi-enclosed spaces. A less contemporary, boxy monitor and a small speaker are installed inside a cubby, flush with the wall, at around torso-level. This monitor plays video scenes at a faster pace, with graphics and/or text occasionally superimposed over the video: game-like graphics or icons and text like "girl on" and "mystery meat being scanned" that give the sense of someone playing with or manipulating the video. These scenes seem more distantly familiar. People in the scenes are going about their business in familiar-seeming public and semi-private spaces, but keep watching: people, some dressed, some nude, appear to be 'caught on tape' in everyday acts, such as working at a desk in an office, or dramatic acts, such as fighting in hallways. These videos have a grainier, washed-out appearance, and the clothing and hairstyles worn by people in the scenes suggest an earlier time. Pausing here, you catch strong ambient noise coming from the museum space immediately outside the entryway, such as footsteps, doors closing, talking, and you catch the occasional sound of a quiet voice coming from the small speaker positioned above the monitor. More distinctly you catch phrases spoken by a woman, the sound coming from around the corner: "Swing your lantern higher." Some information about the artwork is printed on the wall to your left: the name of the artist and collaborators, the title and date of the work, how it came to the museum, and a few sentences introducing the work.

Spin to your right, and follow other *Predictive Engineering*[3] visitors out of the entry passageway, around the corner and toward the large center room. Pause and look up. Placed on a high red shelf in a corner just outside the entry, an older, boxier monitor is angled downwards, towards you, for easier viewing. Video scenes of similar everyday and dramatic acts are playing at a fast pace here as well, with graphics and text appearing often and moving over the video: "Alert: collective security accountability now on." In the opposite corner behind you, another high red shelf is angled downwards. Loose black cables dangle from a wall socket close to the ceiling and fall around this shelf, but the shelf is empty. Was a monitor ever on this shelf?

Go into the large center room. This room's rectangularity feels expansive, like a small studio, and is made active by precisely positioned props, and the playing of video and audio, and by visitors moving around the room. The room's two end walls, near the entrance and exit passageways, are covered by quadrants of large mirrors. Mirrored reflections add a sense of movement or journeying to the space. The supremely high ceiling is painted black, and the walls are white. The lighting is subdued, but interrupted by spotlights softly punctuating the middle course of the room, with one warm, red-hued spotlight trained on the space surrounding a microphone attached to a boom arm pointed

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

inwardly into the room and positioned at face-level on one of longest walls. Two visitors have stopped to try the microphone by speaking and whistling into it.

Move around this rectangular space. Four large flat panel display monitors edged with yellow and black hazard tape are attached to two lengthy trusses in the middle of the room, hanging several feet from the high ceiling. Like the first flat panel display monitor positioned outside the gallery, these four monitors show new scenes moving at a mellow, rhythmic pace. The monitors are positioned on the trusses in ways that make it easy for you to see what is showing on more than one screen at a time. Abundant lengths of licorice black cables fall from the high ceiling and are coiled or draped behind the monitors and around the trusses. Several cables are plugged into nothing: they are dead-ended. You spot a small white drone and four striped balls placed incongruously on the trusses.

The cable-draped trusses and the scenes playing on the four monitors are the most attention-getting visual elements in this room. Visitors around you pause to watch the monitors. The videos cut repeatedly to show familiar scenes and familiar people: it is you and the other visitors in the scene, and spaces that you believe must be the museum. The same scenes can appear simultaneously on more than one screen. You and the other visitors are being watched and recorded by a surveillance camera. These are your scenes. Some visitors begin taking pictures of the monitors when they see themselves appear in the scenes, their faces tipped upwards, towards the monitors, given the height of the surveillance camera attached to the truss.

The video scenes played here are interrupted occasionally and for a beat by the message seen earlier: "Alert Predictive Engineering^3 is currently repairing Predictive Engineering and Predictive Engineering^3." Periodically, you hear very clearly a woman's voice directing curious announcements or pronouncements, and advisories or advice to you, as if broadcast from an unseen control center. She speaks in a smooth, low voice: "Popular control is not happiness." Intermittently, a message is delivered more urgently or, you hear it at a greater volume: "Reach higher, go deeper. Reach higher, go deeper." You realize that the voice messages, the changing video scenes, and the superimposed graphics can be triggered when you or another visitor passes near one of the red-lit sensors placed discretely and evenly at about knee-height along the longest walls in this room. The sensors flash red-green-blue when someone approaches. Voice: "There's a metronome between the two screens."

You and other *Predictive Engineering*^{3} visitors spend more time circulating around this center room. The room's design presents a sense of watchfulness and remove from what is happening outside the room, in its narrowness and

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

openness, the surveillance scenes playing on the large flat panel display monitors, the spoken messages, the mirrors, and the dim lighting. There is an adventurous, but not suspenseful, sensation in being in this room: the rhythmically paced scenes of familiar surroundings and people draw you into watching from different positions, and the messages sounding-out from different speakers positioned around the room encourage you to wander in exploration, since there are no seats. The room and the props may be familiar. It is like walking into a high-control area, where you are shown, or begin to know that you are being watched and recorded. Voice: "The invisibility committee will meet you now."

You make a move toward the exit passageway. This is the other semi-enclosed space flanking the center room. Here, another older, boxier monitor is positioned high on an angled red shelf, with black cables plugged into the ceiling, showing at a fast pace what you imagine or believe are similar or the same scenes seen on the two older model monitors at the entry passageway. You are reminded, as you exit *Predictive Engineering*^{3}, of where you have just been.

Aside from any pictures you took when you spotted yourself in the scenes in the center room, you can take something of *Predictive Engineering*^{3} out of the museum with you. At the entry and exit passageways, an invitation is made to subscribe to a text service called *Security by Julia*. Subscribing to the service will get you a once-daily text message sent at noon by the artist via an application script, much like the alerts received when moving through *Predictive Engineering*^{3}. The text from June 18, 2016 read "Maintain your rotations, please, maintain your rotations in this area."

References

Artsy. 2015. "10 Artists Who Defied Classification in 2015." www.artsy.net/article/the-art-genome-project-10-artists-who-defied-classification-in-2015.

Baca, Murtha, and Patricia Harping, eds. 2014. "Categories for the Description of Works of Art." http://www.getty.edu/research/publications/electronic_publications/cdwa/.

Baca, Murtha, and Visual Resources Association. 2006. *Cataloging Cultural Objects: A Guide to Describing Cultural Works and Their Images*. Chicago: American Library Association. http://cco.vrafoundation.org/.

Bloch, Judy and Suzanne Stein, eds. 2016. *San Francisco Museum of Modern Art 360: Views on the Collection*. San Francisco: San Francisco Museum of Modern Art.

Grau, Oliver, head. Media Art Research Thesaurus http://mediaartresearch.org/search/thesaurus-hierarchical.html

Freeman, Julie, Geraint Wiggins, Gavin Starks, and Mark Sandler. 2015. "A Concise Taxonomy for Describing Data as an Art Material." Proceedings of the IEEE VIS 2015 Arts Program,

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM

VISAP'15, Chicago, Illinois, October 25th–30th.
http://visap.uic.edu/2015/VISAP15-Papers/visap2015_Freeman_ConciseTaxonomy.pdf.
International Committee for Documentation (CIDOC). 2010. "What's in LIDO?"
http://network.icom.museum/cidoc/working-groups/lido/what-is-lido/.
International Council of Museums (ICOM). "Object ID."archives.icom.museum/object-id/checklist.html International Council of Museums
https://archives.icom.museum/object-id/checklist.html.
Lozano-Hemmer, Rafael. 2015. "Best Practices for Conservation of Media Art from an Artist's Perspective."
https://github.com/antimodular/Best-practices-for-conservation-of-media-art.
Older, Jules. 2016. "Scher at SFMOMA." www.youtube.com/watch?v=wH7fRLYWHrw.
Pepi, Mike. 2014. "Is a Museum a Database?: Institutional Conditions in Net Utopia." *e-flux* 60 (December).
www.e-flux.com/journal/is-a-museum-a-database-institutional-conditions-in-net-utopia.
Rinehart, Richard, and Jon Ippolito. 2014. *Re-Collection: Art, New Media, and Social Memory*. Cambridge: MIT Press.
Scher, Julia. 2016. "Predictive Engineering."
https://www.sfmoma.org/watch/julia-scher-predictive-engineering/.

Brought to you by | The National Library of the Philippines
Authenticated
Download Date | 10/11/19 4:08 AM