PROPRIETARY AND OPEN SOURCE COLLECTIONS MANAGEMENT SYSTEMS: TWO CASE STUDIES

by

Niloufar Ameli

Bachelor of Fine Arts York University Toronto, Ontario 2008

A thesis presented to Ryerson University

in partial fulfillment of the requirements for the degree of Master of Arts

in the Program of Film and Photography Preservation and Collections Management

Toronto, Ontario, Canada, 2015 ©Niloufar Ameli 2015

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my thesis may be made electronically available to the public.

Abstract

Proprietary and Open Source Collections Management Systems: Two Case Studies Master of Arts, 2015
Niloufar Ameli
Film and Photography Preservation and Collections Management
Ryerson University

This thesis is comprised of two case studies on the selection and implementation of systems Mimsy XG at Toronto International Film Festival (TIFF) and AtoM and Islandora at Ryerson University's Archives and Special Collections. These studies highlight the changes cultural heritage institutions are currently undergoing to digitally structure, manage, and access collections data by using a relational collections management system (CMS). The findings show that the success of a CMS is largely determined by the type of cultural heritage institution implementing it (archives, libraries, and museums), the types of collections within that organisation, in addition to institutional mandates and requirements. The ways in which a CMS platform answers these requirements through the data model and proprietary or open-source implementations determine the outcome of an institutional transition to a CMS.

Acknowledgments

First I would like to thank my family and friends for their invaluable support.

My sincere gratitude goes to Melissa Neil, Curtis Sassur, and M.J. Suhonos, without whom this research would not have been possible.

A thank you to Dee Psaila for encouraging me to find a research topic I would see myself exploring for years to come.

Lastly, I would like to thank my thesis advisors and readers David Harris, Marta Braun, and Julienne Pascoe for their continued patience and guidance.

Table of Contents

I.	Introduction	1
II.	Literature Review	3
III.	Cultural Heritage Model	10
IV.	Data Model	15
V.	Proprietary vs. Open Source	21
VI.	Research and Analysis	24
M	Aethodology	24
F	Findings	26
	1. TIFF's Move to Mimsy XG	26
	2. Ryerson University's Move to AtoM and Islandora	33
VII.	. Conclusion	39
App	43	
Bibl	57	

List of Appendices

Screenshots of Mimsy at TIFF	43
Fig. 1.1 – An object record in the Library View of the Objects Authority	43
Fig. 1.2 – An object record in the Archives View of the Objects Authority	44
Fig. 1.3 – An exhibition record in Exhibitions View of the Exhibitions Authority	45
Screenshots of AtoM at Ryerson University's Archives and Special Collections	46
Fig. 2.1 – Object record entry page with contracted drop-down fields	46
Fig. 2.2 - Object record entry page with expanded drop-down fields	47
Fig. 2.3 – Object record entry page cont'd.	47
Fig. 2.4 – Object record entry page cont'd.	48
Fig. 2.5 – Object record entry page cont'd.	48
Fig. 2.6 – Object record entry page cont'd.	49
Fig. 2.7 – Object record entry page cont'd.	49
Fig. 2.8 – Object record entry page cont'd.	50
Fig. 2.9 – Object record entry page cont'd.	50
Fig. 2.10 – Object record entry page cont'd.	51
Fig. 2.11 – Object record entry page cont'd.	51
Fig. 2.12 – Object record entry page cont'd.	52
Fig. 2.13 – Object record entry page cont'd.	52
Fig. 2.14 – Object record entry page cont'd.	53
Fig. 2.15 – Object record entry page cont'd.	53
Fig. 3.1 – An object record in AtoM	54
Fig. 3.2 – Object record cont'd.	54
Screenshots of Islandora at RULA	55
Fig. 4.1 – Object record entry page for Theses and Dissertations with expanded drop-down fields	55
Fig. 4.2 – Object record entry page cont'd.	55
Fig. 5.1 – An object record in Islandora	56
Fig. 5.2 – Object record cont'd	56

I. Introduction

The success of a collections management system (CMS) depends on the cultural heritage context and the data model chosen to address the needs and requirements of the institution in which it is used. In cultural heritage institutions, CMS selection and implementation is further complicated by the presence of different types of collections—special collection, museum, library, archival—within one organisation as well as competing management needs, such as access and preservation. Additionally, proprietary vs. open source CMS options must be weighed against institutional mandates and available resources.

A CMS is defined as a computer system that assists institutions with the management of collection information. A CMS is software that provides a means for cataloguing entry, search, retrieval, preservation, and access. It is a digital means to avoid duplication and other collections-management concerns that stem from managing sets of metadata. CMSs are designed with a back and front end to accommodate functional requirements of internal (collections staff) and increasingly external (the public) users. A CMS is ultimately a database system, "an organised collection of logically inter-connected data items" based on an underlying data model created to represent the needs of the institution where it is implemented, related to the broader cultural heritage context. ²

CMSs are either proprietary or open-source in model. Proprietary systems are the exclusive property of their developers or publishers. Access to proprietary source code is usually

¹ Annamaria Poma Swank, *Collection Management Systems Report* (report presented at the Fondazione Rinasimento Digitale, Italy, 2008) 2-8.

² Mark Levene and George Loizou, *A Guided Tour of Relational Databases and Beyond* (London: Springer-Verlag, 1999), 1.

restricted and not generally available to the user. Proprietary systems require payment for installation – as well as for further service and maintenance. In contrast, open source systems are those distributed with their source code available to the user to use and modify as he or she sees fit. This type of system is also referred to as Free and Open Source Software (F/OSS).³ The "free" in this instance refers to the freedom to access, use, modify, and redistribute the software and its source code.⁴

Through two case studies, this thesis will examine the underlying cultural heritage and metadata factors that determine the selection and effective implementation of proprietary and open-source collections management systems. This thesis will evaluate the following: the relational data model represented by the systems studied; the cultural heritage framework for implementing a CMS; and the use of proprietary vs. open-source systems to fulfill institutional requirements.

There are many institutions currently undergoing transformations with respect to a digital approach to their collections and the ways in which they structure institutional data. Toronto International Film Festival (TIFF) and Ryerson University are two institutions undergoing such transformative periods in their approach to collections management using digital applications. This research provides a timely study on goals, approaches, implementation, and outcomes for cultural heritage institutions transitioning to a new CMS platform. The findings can be used to assist other institutions considering similar digital methods for collections management.

³ Joseph Feller, Brian Fitzgerald, Scott A. Hissam, and Karim R. Lakhani, eds., *Perspectives on Free and Open Source Software*, (Cambridge: MIT Press, 2005), xvii.

⁴ Steven Weber, *The Success of Open Source* (Cambridge, MA: Harvard University Press, 2004), 5.

II. Literature Review

The following literature review was conducted to locate literature of relevance on data models and database design, the cultural heritage model, as well as the debate on proprietary vs. open-source systems. Given that the development and use of open-source software for managing collections information is relatively new in cultural heritage discourse, there is little documentation or research in this area at this time.

Resources on the Cultural Heritage Model

Books such as Introduction to Metadata, Managing Image Collections: A Practical Guide, and Linked Data for Libraries, Archives and Museums: How to Clean, Link and Publish Your Metadata are excellent resources for understanding the nuances of the cultural heritage model. These texts discuss approaches to metadata management by each type of cultural heritage institution, in addition to metadata standards introduced to bridge gaps between arrangement and description methods in libraries, archives, and museums. Version 3.0 of Introduction to Metadata, edited by Murtha Baca, is a good resource for understanding what metadata is, how it is approached by libraries, archives, and museums, and best practices for metadata creation and maintenance. The electronic version of this book is also freely accessible on the Getty Publications website. Managing Image Collections: A Practical Guide by Margot Note is primarily concerned with photographic collections management issues, including that of photographic collections metadata across cultural heritage institutions. The structure of this

⁵ Tony Gill, Anne J. Gilliland, Maureen Whalen, and Mary S. Woodley, *Introduction to Metadata*, ed. Murtha Baca (Los Angeles: Getty Publications, 2008).

 $http://www.getty.edu/research/publications/electronic_publications/intrometadata/index.html.\\$

⁶ Margot Note, *Managing Image Collections: A Practical Guide* (Oxford: Chandos Publishing, 2011).

book is clearly laid out, and the many tables and figures of organised information are particularly helpful. *Linked Data for Libraries, Archives and Museums: How to Clean, Link and Publish Your Metadata*, by Seth van Hooland and Ruben Verborgh is a manual for cultural heritage institutions changing their approach to metadata. The book defines metadata, describes data modelling, cleaning, reconciling, and publishing for the purposes of linking and making collections data accessible on the web. It is written to aid those with minimal technical expertise to make information management decisions for cultural heritage collections with a particular focus on recent trends in linked data.

Resources on Databases

In order to formulate a better understanding of databases and their design, collections managers can reference the following texts which survey database technologies. *The Database Relational Model* in particular examines relational databases and serves as a good introduction to the evolving nature of software. *The Structure of the Relational Database Model* explains how databases function within the relational model. In the *Guided Tour of Relational Databases and Beyond*, differing models of database design such as the currently accepted standard, the relational model are discussed at length. The introductory sections to the above three texts are especially illuminating, although the technical nature of the overall content can render them inaccessible and overwhelming to the reader without a background in software development.

⁷ Seth van Hooland and Ruben Verborgh, *Linked Data for Libraries, Archives, and Museums: How to Clean, Link and Publish Your Metadata* (Chicago: Neal-Schuman, 2014).

⁸ CJ Date, *The Database Relational Model* (USA: Addison Wesley Longman, Inc., 2001).

⁹ Jan Paradaens, Paul De Bra, Marc Gyssens, and Dirk Van Gucht, Monographs on Theoretical Computer Science, vol. 17, *The Structure of the Relational Database Model* (Berlin: Springer-Verlag, 1989).

¹⁰ Mark Levene and George Loizou. *A Guided Tour of Relational Databases and Beyond*, (London: Springer-Verlag, 1999).

Shannon Leigh Perry's 2006 thesis *So, You Want a Database?: A Beginner's Guide to Database Technology for Small Cultural Institutions* serves as a how-to manual for institutions aiming to understand how databases function in general terms. While this thesis does not provide a cross-comparative study between proprietary and open source systems it does provide the reader with an essential understanding of databases, CMS, and their function.¹¹

Elana C. Carpinone's 2010 thesis *Museum Collections Management Systems: One Size does Not Fit All* takes an in depth look at some widely known proprietary systems utilized by major institutions. Carpinone maps the history of database technology in previous decades, and proceeds to describe and analyse each of the seven proprietary systems examined. The author's analysis presents a lot of quantitative data, illustrating the types of institutions that employ the systems studied and the associated costs.

Resources on Proprietary vs. Open-Source Systems

Proprietary websites such as those of *Mimsy XG* and *Minisis Inc*. function as marketing platforms for the sale of their respective CMSs.¹³ What these sites offer, however, are attribute highlights of each system and vendor marketed. The Minisis website places primary emphasis on the vendor's longstanding history, a list of clients utilising each of the four systems offered, with short lists of what each system can offer for collections management. Some main attributes listed include: standards compliance, hierarchy and description capabilities, customisation, web search

¹¹ Shannon Leigh Perry, *So, You Want a Database?: A Beginner's Guide to Database Technology for Small Cultural Institutions* (MA thesis, Ryerson University, 2006).

¹² Elana C. Carpinone, *Museum Collections Management Systems: One Size does NOT Fit All* (MA thesis, Seton Hall University, 2010).

¹³ *Mimsy XG*, MIMSY, accessed July 14, 2015. http://www.willo.com/mimsy_xg/.; *Minisis Inc*, MINISIS, accessed July 14, 2015. http://www.minisisinc.com/index.html.

ability, and tracking of objects. Using screenshots, the Mimsy website illustrates the way the interface of the system would look, and the features the vendor believes to be the major selling points for the product: its thirty-plus year history, versatility, and ease of use.

The following books examine the particular attributes that define open-source software and set it apart from proprietary software, but also examine its social, political, and economic implications. *Open Sources 2.0: The Continuing Evolution* is a group of essays by contributors respected in the open-source software industry. ¹⁴ This book provides in-depth definitions of what open source is and how it functions in both technical and political terms. *Perspectives on Free and Open Source Software*, similar to that of *Open Sources 2.0*, is a set of edited essays on the topic of open-source software. ¹⁵ Given the high use of technical jargon and mathematical equations, this book is too advanced for the novice reader without a background in software development. *The Success of Open Source* is written by an expert in the field, Steven Weber, a professor of political science and information technology. ¹⁶ In terms of content, the book is accessible to the non-familiar reader and without much technical information. *The Success of Open Source* is primarily concerned with charting the historical evolution of open source while examining how it works, the participants involved, and its larger effects on the political economy.

Jessica Bushey's paper entitled *ICA-AtoM*: *Open-Source Software for Archival Description* is a recent article that charts the rise of ICA-AtoM (International Council on

¹⁴ Chris DiBona, Mark Stone and Danese Cooper, eds., *Open Sources 2.0: The Continuing Evolution*, (Cambridge: O'Reilly Media, 2005).

¹⁵ Joseph Feller, Brian Fitzgerald, Scott A. Hissam, and Karim R. Lakhani, eds., *Perspectives on Free and Open Source Software*, (Cambridge: MIT Press, 2005).

¹⁶ Steven Weber, *The Success of Open Source* (Cambridge, MA: Harvard University Press, 2004).

Archives – Access to Memory) as an open-source alternative to proprietary systems for collections management.¹⁷ AtoM was developed by Artefactual Systems in 2005 under the direction of the International Council on Archives. With its last release in December 2011, ICA-AtoM 1.2, AtoM has since been a subject of praise, as well as the recipient of funding and technical expertise for further development.¹⁸

Bushey outlines how the system is grounded in open-source values of open collaboration, its number of users as of 2012, and the system's significance as demonstrated by three organisations' migration from proprietary systems to AtoM: British Columbia Archival Union List—Memory BC; Archives Association of Ontario—Archeion Network; and the City of Vancouver Archives. This paper is a great alternative resource on an open-source CMS, however it must be noted that at the time of publication Bushey was AtoM product manager at Artefactual Systems. Given that this resource is written on behalf of the CMS vendor, the paper therefore exhibits a bias toward open source development much in the same way as the proprietary websites reviewed previously.

Canadian Heritage Information Network (CHIN)'s *An Introduction to Free and Open*Source Software for Museums provides a concise description of open source, and how it may be of benefit to museums.²⁰ Although the document is not a comparative study on open source and proprietary systems on the market, it acts as a good theoretical introduction to the topic of open

¹⁷ ICA-AtoM: Open Source Archival Description Software, ICA-AtoM, accessed February 1, 2014. https://www.ica-atom.org/download/ICA-AtoM_JBushey.pdf.

¹⁸ *Ibid.*, 1-3.

¹⁹*Ibid.*, 3, 13-16.

²⁰ An Introduction to Free and Open Source Software for Museums, CHIN Professional Exchange, accessed February 1, 2014. http://www.pro.rcip-chin.gc.ca/carrefour-du-savoir-knowledge-exchange/intro_code_source_ouvert-intro_open_source-eng.jsp.

source for institutions considering that route. There is surprising amount of detail in this document, in spite of its brevity. The document's reliability is also validated by the fact that it is produced by the Canadian government with the aim of connecting heritage institutions and providing resources to them for further development.

Documents such as the *Collection Management Systems Report* by Annamaria Poma Swank, *Choose a CMS Software* survey by Collections Trust, and the *Collections Management Software Criteria Checklist* by CHIN are all good resources for collections looking to upgrade to another system or acquire a new one. ²¹ These resources mainly provide collections with descriptions of popular proprietary systems on the market, as well as criteria to be considered by cultural institutions when considering a new CMS. The 2012 checklist by CHIN was created as a way to define collections management requirements at institutions, and is therefore not concerned with the type of system, but rather system features and capabilities that can assist institutions with their collections management practices.

The report and survey, in contrast, cover numerous popular CMSs by comparing what they can offer. *Collection Management Systems Report* by Swank surveys eighteen vendors who offer at least one CMS among other applications.²² All of the vendors listed produce proprietary CMSs and no open-source options are explored. In *Choose a CMS Software*, only one out of the eighteen systems covered is open-source.²³ The report and the survey therefore do not provide a

_

²¹ Annamaria Poma Swank, *Collection Management Systems Report* (Italy: Fondazione Rinasimento Digitale, 2008).; *Choose a CMS*, Collections Trust, accessed June 14, 2015, http://www.collectionstrust.org.uk/choose-acms.; *Collections Management Software Criteria Checklist - 2012*, Canadian Heritage Information Network, accessed June 14, 2015, http://www.rcip-chin.gc.ca/carrefour-du-savoir-knowledge-exchange/criteres_gestion_collections_management_checklist-eng.jsp.

²² Swank, Collection Management Systems Report.

²³ Choose a CMS, Collections Trust.

balanced representation of the range of CMSs currently available for collections management but rather mostly highlight well-known proprietary systems.

The user manual of the AtoM website provides a sense of the system's capabilities.²⁴
However, the website is rich in technical jargon which can make the information presented tough to understand for someone without a background in systems development. The website for open-source system Islandora is relatively simpler in content and is therefore easier to understand.²⁵
Both sites for AtoM and Islandora outline key features, and how each system can be used to manage digital content. A key feature of open-source websites are opportunities for discussion and open advancement of software. The AtoM and Islandora websites for example provide access to documentation, wikis, and forums.²⁶

The system websites referenced in this review are rooted in vested interests, acting as further marketing for each system, while comprehensive academic studies of how proprietary systems function in comparison to open-source counterparts are rare.

_

²⁴ What is AtoM?, AtoM, accessed July 14, 2015, https://www.accesstomemory.org/en/docs/2.2/user-manual/overview/intro/.

²⁵ Why are Organisations Migrating to Islandora from Other Digital Repository Systems?, Discoverygarden: Managing Digital Content, accessed July 14, 2015, http://www.discoverygarden.ca/islandora-overview/.

²⁶ Community: Get Support, AtoM, accessed August 28, 2015, https://www.accesstomemory.org/en/community/support/.; Islandora Community, Islandora, accessed August 28, 2015, https://wiki.duraspace.org/display/ISLANDORA715/Islandora+Community.

III. Cultural Heritage Model

Decisions that determine and implement a successful transition to a CMS platform must consider the underlying cultural heritage context. Libraries, archives, and museums are types of cultural heritage institutions which aim to preserve and create better access to cultural artifacts. Each type differs from one another in its institutional requirements as well as approaches to data management. These cultural heritage contexts inform the institutional requirements and thus the underlying data models and CMS approaches.

Libraries are service-oriented institutions dedicated to providing access to "published, non-unique items, such as books and serials, which are produced in multiple copies." As a result, libraries catalogue at the item-level, focusing on subjects, titles, and authors. The organisation of resources is subject-based, with access and retrieval based on classification systems such as Library of Congress call numbers. Given that library metadata includes indexes, abstracts, and bibliographic records created with data structure (MARC21), content (AACR2), and value standards (Library of Congress Authorities) in mind libraries require databases and repositories that meet these standards in order to manage their collections. ²⁸

Archives identify and preserve "the evidential value of records and artifacts in and over time" and they assist "researchers with their analysis and interpretation" of objects such as "unpublished materials and primary sources." Provenance and original order (respect des fonds) are the two key guiding principles of archives. These two principles govern the structure

²⁷ Note, *Managing Image Collections*, 109.

²⁸ Gill et al, *Introduction to Metadata*, under "Setting the Stage,"

http://www.getty.edu/research/publications/electronic_publications/intrometadata/setting.html.

²⁹ *Ibid.*; Note, Managing Image Collections, 109.

of fonds and that of archives by ensuring that fonds are kept separate based on their source, and that the materials are kept together according to the structure of their use and archival context.

Archival metadata "includes accession records, finding aids, and catalog records" and is in the form of hierarchical relationships exhibiting many levels –fonds, series, files, and item. In order to make records accessible, archives describe the materials, summarise the context and content at each level, and note any access restrictions for the materials.³⁰ Archives require databases and repositories suited to archival data structure (EAD) and data content standards following the guiding principles of archives (DACS in the US, RAD in Canada) to manage their collections.³¹

Museums on the other hand, showcase unique holdings of art and artifacts in the form of exhibitions and other events. Museums describe these holdings at the collection and the item level. Objects may be grouped by categories such as maker, media, provenance, or historical period and the description may include other details about the object and any associated preservation concerns. The subjective formulation of object attributes and descriptors by museum collections staff, as well as the diversity of museums and their collections can entail varying practices from one museum to the next. Nevertheless, museums have increasingly begun to apply data structure, content, and value standards to describe their collections and can utilise a variety of systems that accommodate the above standards in managing their collections. The Getty Institute has been instrumental in bridging collections information in museums and overcoming idiosyncratic cataloguing practices by creating data structure (CDWA)

_

³⁰ Gill et al, *Introduction to Metadata..*; Note, *Managing Image Collections*, 109.

³¹ Gill et al, *Introduction to Metadata*.

³² Note, Managing Image Collections, 109.

³³ Gill et al, *Introduction to Metadata*.

Lite), data content (CDWA) and data value standards (AAT, ULAN, TGN) now widely implemented and used in museum CMSs.

TIFF

TIFF is a not-for-profit organisation that grew from the ten-day annual film festival in Toronto to film-related programming throughout the year. TIFF's mandate is to present the best of international and Canadian cinema to film lovers and to "transform the way people see the world, through film."

Aside from film screenings, lectures, discussions, festivals, and workshops, TIFF also features a comprehensive Film Reference Library (FRL) housing "the world's most extensive English language collection of Canadian film-related material." The FRL collections include 300,000 production stills and photographic images. In addition, TIFF showcases film-related exhibitions and moving image art to "explore the intersections between cinema and the visual arts." Thus, through its programming and services TIFF caters to the public, filmmakers, scholars, students, curators and museum professionals, offering opportunities for entertainment, learning, and collaboration.

TIFF collections are overseen by the FRL and the Exhibitions departments and comprised of library, archives, and special collections. The two departments began using CMS Mimsy XG in 2014 to manage their varied collections. Library items are catalogued using data content standard RAD, data value standard LCSH for printed matter, and FIAF subject headings for

³⁴Our Story, TIFF, accessed May 15, 2015, http://tiff.net/explore/about.

³⁵ Ibid

³⁶ Film Reference Library FAQ, TIFF, accessed November 24, 2013, http://tiff.net/filmreferencelibrary/library/faq.

audio-visual materials. Special collections items, as well as any loan exhibition items are catalogued using RAD data content standards. The iteration of Mimsy implemented at TIFF has been customised to accommodate the different standards required to maintain these collections through the setup of separate but linked Library, Archive, and Exhibition 'Views' for the FRL and Exhibitions departments to work in.

TIFF is launching Mweb in fall 2015, its first web-based and remotely accessible catalogue search for all TIFF collections entered in Mimsy.

Ryerson Archives and Special Collections

Ryerson University Library's Archives and Special Collections departments operate on the mandate to "acquire, preserve, and make accessible all non-circulating library collections materials," which include university records, archival fonds, curated collections, and rare books. In the Archives department, the aim is to serve the academic institution's heritage through the collection, preservation, and the making of university records and documents accessible to the Ryerson community. The Special Collections department was "established to help support the learning and teaching needs and facilitate the scholarly, research and creative activities of the Ryerson community by acquiring and preserving photography, film and cultural history objects. Therefore, Ryerson Library's Archives and Special Collections cater primarily to students, staff, and researchers interested in Ryerson's history from 1948 to the present, as well as the institution's historical roots, from 1852-1948.

_

³⁸ A&SC Collections Policy information for Potential Donors. Archives and Special Collections, accessed May 15, 2015. http://library.ryerson.ca/asc/donations/.

³⁹ Archives, Archives and Special Collections, accessed May 15, 2015. http://library.ryerson.ca/asc/archives/.

⁴⁰ Special Collections, Archives and Special Collections, accessed May 15, 2015. http://library.ryerson.ca/asc/sc/.

Ryerson Library's Archives and Special Collections utilise RAD as their data content standard for archival description and in August 2015 the departments began to use the web-based and open-source system AtoM to manage their collections. The Archives and Special Collections can be searched remotely via links to the Archives and Special Collections web pages. Ryerson University Library and Archives (RULA) also implemented Institutional Repository (IR) Islandora for managing library-generated digital content in addition to managing the collection of theses, dissertations, and research papers in its holdings.

IV. Data Model

Data models are defined as abstract models or representations of data created for any domain, regardless of the information system used. These models are able to express the requirements of a given domain through tables of entities and their relationships. When designing a database, modelling of data ensures that the needs of an organisation are represented and met by the system being designed, since a data model determines the structure of the database being created, pinpointing how data is stored, managed, and manipulated. 42

Metadata is defined as "structured data about data, information that facilitates the management and use of other information." In other words, metadata "provide users with a standardized means for intellectual access to holdings." Metadata standards are conventions to facilitate the "transfer of information between hardware and software platforms." These standards are especially important for ensuring interoperability as technologies continue to evolve. Metadata standards can be broken down to those overseeing data structure, data value, data content, and data formats. Thus, data models and metadata are significant for managing cultural heritage collections as they provide standard mechanisms for the description of a variety of materials in holdings used by the broader cultural heritage model, and the creation of access to these holdings through the use of software.

.

⁴¹ Gill et al, *Introduction to Metadata*, under "Glossary,"

http://www.getty.edu/research/publications/electronic_publications/intrometadata/glossary.html.

⁴² Julienne Pascoe, *Introduction to Relational Databases* (Lecture Slides, Ryerson University, Toronto, ON, February 2, 2015).

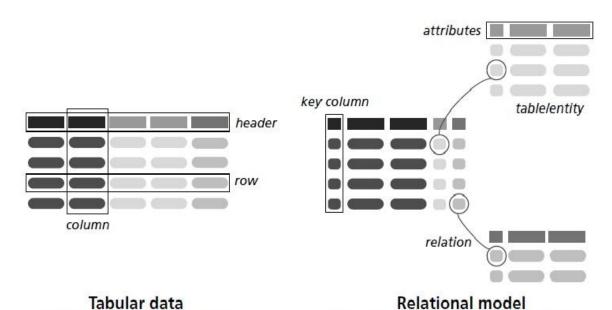
⁴³ Note, *Managing Image Collections*, 117.

⁴⁴ Ibid.

⁴⁵ Ibid

⁴⁶ Gill et al, *Introduction to Metadata*, under "Setting the Stage."

The most widely used and implemented data model, the relational data, was introduced in the 1970s as a way to address the limitations of the previous data model, tabular data. The tabular data model is comprised of data expressed in the form of columns and rows. In this model, data in a given cell gains meaning through the intersection of a column with a given row. The below diagram illustrates the overall structure of a tabular data model vs. that of a relational model.



Each data item is structured as a line of field values. Fields are the same for all items; a header

line can indicate their name.

Data are structured as tables, each of which has its own set of attributes. Records in one table can relate to others by referencing their key column.

48

As in the above diagram, data in the relational model is organised in tables/entities. Each table or entity is defined by its own row of attributes and a key column. Each record in a table

⁴⁷ van Hooland and Verborgh, *Linked Data for Libraries, Archives, and Museums*, 15-24.

⁴⁸ van Hooland and Verborgh, *Linked Data for Libraries*, *Archives, and Museums*, 16, fig. 2.1 detail.

has its own unique key which is used to refer to other tables in what is called a 'relation.' The below diagram provides a closer look at the way data is structured in the relational data model.

	Creator							
	ID	first	st name surname blo Picasso		birth ye	ear death year		
	(43)	Pablo			1881	1973		
	57	Jeff		Koons	1955	null		
	\	\						
			V	/ork				
ID	title	1	creator	collection	year	style		
5	Guernica	Y	43)	20	1937	cubism		
7	First Commun	nmunion 43		(22)	1895	realism		
16	Puppy		57	18	1992	conceptual		
		38		/				
			/	Collection				
			ID/	name		address		
			18	18 Guggenheim		Bilbao		
			20	Museo Reina Sofia		Madrid		
			22	Museo Picasso		Barcelona		

49

In the above diagram data is organised in three tables or entities titled Creator, Work, and Collection. Each table/entity has a row of attributes listed at the top, as well as an ID column for the unique keys associated with the individual records in each table. Records 22 in the Collection table (Museo Picasso) and record 43 (Pablo Picasso) in the Creator table are referenced in the Work table by way of their IDs.

⁴⁹ van Hooland and Verborgh, *Linked Data for Libraries, Archives, and Museums*, 24, fig. 2.2.

All three systems studied in this thesis are based on the current data standard: the relational model. In a document created by the vendor Selago Design to highlight the features of Mimsy for cultural heritage institutions, the CMS is marketed as a system with the following attributes: versatile in managing diverse collection types such as museums, libraries, and archives, customisable in configuration to suit a collection's needs, quick and efficient data entry when cataloguing, capable of generating reports in a variety of formats, web-enabled to share any chosen collections data online, and created around the Oracle database to ensure continued access to inputted data for future access. 50 The system uses modules (Authorities and Activities) to manage diverse museum requirements such as cataloguing, exhibitions, tracking, media, etc. created to suit TIFF's institutional data management needs. The Appendices screenshots of Mimsy at TIFF (figs. 1.1-1.2) show two object records entered into the Object Authority and fig. 1.3 shows an exhibition record in the Exhibition Authority. Module names corresponding to each Authority or Activity in Mimsy are shown at the top of the white cataloguing area. TIFF's Archive, Library, and Exhibitions Views can also be seen in the figs. 1.1-1.3, indicated by the View Selector box at the far top right. Any record links to other Activities or Authorities can be seen in the lower left hand corner or the Link Ledger. Any multimedia files linked to the record are shown in the Media Panel above the Link Ledger.

The user manual of the AtoM website describes the archival description software as possessing the following features: it is a web-based application, its source code is available for use and modification without an initial cost, it is built around the descriptive standards set forth by the International Council on Archives, it can operate in multiple human languages, and it can

-

⁵⁰ *Portfolio: Mimsy XG*, Selago Design, Inc., accessed July 14, 2015. www.selagodesign.com/portfolio/mimsyxg/index.php.

also function as a multi-repository system, for instance when a network of institutions use the same system. The AtoM user manual also lists seven entity types (accession record, archival descriptions, authority record, archival institutions, functions, rights records, and terms) that represent the main data tables required by an archival institution to manage collections information. Secreenshots of Ryerson University's iteration of AtoM are included in the Appendices. Fig. 2.1 shows that drop-down fields are contracted by default and can be expanded by clicking into a field when cataloguing (figs. 2.2-2.15). Additionally, any required fields are marked by asterisks. Figs. 3.1-3.2 show how an object record in the system looks. In fig. 3.1 an Export button can be seen just above the Related Subjects area on the right hand side. This button allows for exporting of records in the main archival standards, EAD and Dublin Core.

The Discovery Garden website responsible for leading the development of the IR Islandora describes the system's features as: configurable to the needs of the user institution, able to search quickly and efficiently, and flexible installation – can be installed on the user institution's server or hosted by Islandora. Additionally, the website emphasises that the project is rooted in industry best-practices and built on well-known open-source systems Fedora and Drupal. It is important to note that while Islandora can be utilised as a CMS, its designing feature is to manage and make accessible digital scholarly collections and materials as an IR.

5

⁵¹ What is AtoM?, AtoM, accessed July 14, 2015. https://www.accesstomemory.org/en/docs/2.2/user-manual/overview/intro/.

⁵² Entity Types, AtoM, accessed August 28, 2015, https://www.accesstomemory.org/en/docs/2.3/user-manual/overview/entity-types/#entity-types.

⁵³ Islandora was originally developed by University of Prince Edward Island's Robertson Library, but is now widely available for implementation at any institution.

⁵⁴ Why are Organisations Migrating to Islandora from Other Digital Repository Systems?, Discoverygarden: Managing Digital Content, accessed July 14, 2015. http://www.discoverygarden.ca/islandora-overview/.

page for theses and dissertations can be seen in figs. 4.1-4.2. These images show the available fields when cataloguing a thesis or dissertation, with the required fields marked by an asterisk. Figs. 5.1-5.2 show the appearance of an object record for a thesis within Islandora.

V. Proprietary vs. Open Source

Proprietary and open-source business practices are representative of the ways software development has evolved over the years within the world markets. In economic terms, proprietary and open-source business practices are indicative of differing understandings of property rights, copyright in particular. The proprietary economic model is concerned with property rights resting solely with that of the property owner where no person other than the owner is permitted to distribute and modify the property in any way. On the other hand, the open source model is interested in the freedom to distribute that property, and allow others to freely modify it for the betterment of society. ⁵⁵

The advantages to implementing a proprietary system for collections management are that the products are better known due to proprietary dominance in the market; that they have been around for a longer period than open-source systems; that they are off-the-shelf products readily available for collections management purposes; and that they provide technical support which results in a reduction of resources required in terms of technical expertise. Disadvantages to implementing proprietary systems include the high purchase price and maintenance fees; service contracts overseeing upgrades and technical changes; costs associated with additional features; limited technical support at a cost; and closed source code due to the proprietary nature of the software.

Advantages to implementing an open source system include lower costs of implementation and system updates; with some systems in fact free in monetary terms; the

⁵⁵ Weber, *The Success of Open Source*, 227.

ability to evolve and adapt the software to the needs of a particular organisation; technical support that is readily available 24 hours a day through collaborative networks of developers and information professionals; and source code that is available for manipulation and change. The development of open-source systems is inherently tied to the development of the open-source community that works together to advance better long-term software development that evolves with the diverse needs of the community it serves. ⁵⁶ Disadvantages to open-source systems are that they are newer to the marketplace and software landscape and therefore lesser known, and that development, implementation, and management of open-source systems requires more resources, including staff training and IT support.

While a majority of institutions still employ proprietary CMSs to manage collections, open-source systems exhibit the potential to eventually overtake them.⁵⁷ Two of the three sources interviewed for the purposes of this thesis commented on the rise of open-source systems and discussed the compatibility and openness of source code as open-source's winning attributes.⁵⁸ The continued rise of open source and its greater availability would mean an increase in more flexible, cost-effective CMS options available for cultural heritage institutions and more choices in systems as a whole.

Nevertheless, the ultimate preference of proprietary vs. open source is at the discretion of the institution. Each type of cultural heritage institution has diverse requirements that are met by each model. Proprietary systems appear to be the most popular choice for cultural heritage institutions such as museums since many such organisations lack the funding and staffing to

_

⁵⁶ Weber, The Success of Open Source, 227.

⁵⁷ Swank, Collection Management Systems Report, 24-26.

⁵⁸ Curtis Sassur and M.J. Suhonos, interview by author, Toronto, February 28, 2014.

develop their own open-source software. Whereas larger academic institutions (such as libraries) are able to afford in-house IT teams to develop open-source solutions for their digital collections. The changing open-source landscape means greater flexibility in the implementation and management of open-source systems: AtoM for example offers institutional support to those institutions that do not have the IT staffing.

VI. Research and Analysis

Methodology

This component of the thesis comprises of two case studies undertaken to determine how and why two institutions chose proprietary and open source CMSs. The two institutions that agreed to my request to interview their staff were TIFF and RULA. The research was based on interviews with Melissa Neil, former Registrar, Mimsy Systems Administrator and Project Manager at TIFF and Curtis Sassur, Coordinator of Ryerson University's Archives and Special Collections, along with M.J. Suhonos, Digital Technologies Development Librarian at RULA. These two institutions' differing choices of CMSs are noteworthy as they can tell us about current collections management options in managing collection data.

At TIFF I interviewed Melissa Neil about her understanding of the process and involvement with the migration to Mimsy from the previous system Star, in place at the FRL. At RULA I interviewed Curtis Sassur and M.J. Suhonos, about the ongoing move from Minisis to two open source counterparts: AtoM and Islandora.

I asked each interviewee to describe why their institution selected its new system, how they undertook its development, and the institution expected outcome of implementing that particular CMS to be. My interviews attempted to get a better understanding of each institutions' values and requirements. The interviewees described perceptions surrounding the CMS they

previously utilised such as legacy systems Star at TIFF and that of Minisis at Ryerson University's Archives and Special Collections.⁵⁹

The interviews were completed over a period of four weeks, in the months of March and April of 2014. Upon completion of the interviews, I transcribed each segment and created follow-up questions for the interviewees for better clarification.

What I set out to accomplish in these interviews was to get an understanding of given institution selects and implements a CMS based on the system's capabilities to meet the requirements of the institution. The data that emerged demonstrated there are many ways to facilitate access to collections using proprietary and open source systems currently available but that the ultimate choice was determined by the needs of a given type of cultural heritage institution and its kinds of collections.

⁻

⁵⁹ A legacy system is defined as a previous or outdated computer system.

Findings

1. TIFF's Move to Mimsy XG

Melissa Neil was the Registrar, Mimsy Project Manager and Systems Administrator, when I interviewed her in March 2014 to gain an understanding of what Mimsy has offered TIFF. The interview centred on the impetus to adopt a CMS, why TIFF chose Mimsy, what the project of selection was, and what the implementation of the system looked like, as well as any effects it had on the institution's collections management procedures. Since Neil was no longer in the position of Registrar and Project Manager in 2015, follow-up emails were sent to the new acting Project Manager Michelle Lovegrove Thomson in May and June of 2015 to better understand where the project stood and what if any of the expected outcomes had occurred.

Neil described the process as having begun in 2013 when she was hired to "project manage the implementation of Mimsy" at TIFF.⁶⁰ Neil explained that the search for a new CMS at TIFF began a year earlier. This was to allow time for research on available CMSs, the submission of a business case to the organisation itself to acquire funding for the project, and the submission of a Request for Proposal (RFP) to rank vendors. The RFPs functioned as a way to narrow down the institution's choices from a longer list of CMSs generated by a previous staff member before Neil came on board.

According to Neil, the impetus for investing in Mimsy came from the Exhibitions department's programming:

[TIFF] wanted to tour the Cronenberg exhibition. And there was a plan to potentially keep doing that, to keep generating exhibitions and putting them out in the world as a

⁶⁰ Melissa Neil, interview by author, Toronto, March 24, 2014.

possible revenue generator drawing on our collections and loans. So there was a push from the Exhibitions department which said if you want us to do this, we need the tools to do it: we need a collections management system.

Neil talked about how while the FRL was an older department that existed when the organisation was located on Carlton Street prior to the move to King West, the Exhibitions department only opened in 2010 coinciding with the opening of the TIFF Bell Lightbox. Neil stated that unlike the FRL, Exhibitions had never employed a CMS since opening. Exhibitions, she added, "didn't have as many collections as the Library [FRL] did to justify the expense, until the level of exhibitions activity reached what it did." Neil described that, due to the close-knit relationship between the FRL and Exhibitions departments, once Mimsy was utilised for Exhibitions the discussion then shifted to: "if you're going to [employ a CMS for Exhibitions], what about the library collections [FRL]."

The FRL at the time employed a system by the proprietary vendor Cuadra, named Star, and the relationship was over twenty years old. Neil discussed how the FRL had not implemented any system upgrades for Star over the years and while functional, the system was not as efficient as it would have been with regular upgrades and maintenance. Additionally, Star was described by Neil as a non-relational system which would make it a predecessor to the relational systems of today. 61

In describing the effects of not applying any system upgrades to Star for more ten years,
Neil was quick to add "...at least they had a system. And it worked for them. For those who were

⁶¹ A Guided Tour of Relational Databases and Beyond by Mark Levene and George Loizou explains that a relational database is one modelled after the relational model introduced by E.F. Codd in 1970, a model which has become the dominant one not only in database research, but also in commercial terms in the database management system market. *The Structure of the Relational Database Model* by Jan Paradaens, Paul De Bra, Marc Gyssens, and Dirk Van Gucht further explains how in a relational model unlike its other contemporary models—information is stored by way of tables, referred to as relations. This model is explained in further detail in the Data Model section.

working with the data every day, and knew how to make the system work, they were pretty comfortable with it."

TIFF's special collections data on the other hand was never stored in any system. "Each collection had a finding aid and that probably would be the extent of the information captured about the collection, other than in the initial intake and acquisition documents. These were kept in Word files, Excel documents, and PDFs: which meant they were completely unsearchable, other than opening up a binder or looking through files one by one, on the server."

Neil stated how the lack of information captured upon acquisition, in other words the lack of cataloguing at the fonds level and "not actually going down to the item level," as well as the lack of searchability of special-collection finding aids meant that over time TIFF was faced with the problem of not knowing just how many items were in any given special collection. This problem would be resolved in the process of cleaning up the data in the finding aids and migrating that data into Mimsy. TIFF was then able to able to assess just how many items were in each collection.

One reason TIFF chose Mimsy as its CMS was the flexibility to configure it as the institution saw fit: "Mimsy was highly configurable and we thought that what [specs] we didn't see, we could easily customize to our needs."

Other reasons for the choice, Neil suggested, were Mimsy's ability to capture large amounts of information and TIFF's aspiration to adhere to museum standards:

I think that there was a tremendous amount of options for activities, lot of authorities, and we felt that there was an enormous amount of potential to capture as much information we may need. I think that at the Exhibition-level we knew that we needed to have many

objects in containers, and be able to move objects around the world, and have a container leave a crate, and we felt really confident that [Mimsy] was going to be able to manage all that information for us. I think it was important for this institution to feel like it is operating at a museum standard. That's always the aim here: to be operating and exceeding international museum standards. It's an aspirational goal and it was really important for [TIFF senior staff] to see that all these big name museums using the system.

TIFF narrowed down choices for CMSs down to three by extensive research prior to the RFP. The three chosen were then contacted for proposal submissions: Adlib, Mimsy, and TMS were the top three systems in the running. Once TIFF received the proposals, TIFF asked each CMS vendor for a demo. Neil discussed how TIFF utilised the rating matrix below inspired by the CHIN checklist, in order to rank the three systems out of a total score of hundred.⁶²

							Notes	
Fit with			Flexibility /	Technical	Relevant			
Requirements	Lifecycle Cost*	Support	Scalability	Architecture	Experience	Sum	Pros	Cons
Score out of 30	Score out of 20	Score out of 15	Score out of 15	Score out of 10	Score out of 10			

Neil stated that representatives from Exhibitions and FRL were asked to rate the three systems. Three diverse items that were previously catalogued were used as test objects under museum standards, archival standards, and library standards with each CMS being rated to see which performed the best for TIFF's collections management purposes. At the end of this process, the top two CMSs were that of Mimsy with a total score of 93, and Adlib with a score of 92. Neil explained that although both Adlib and Mimsy had received close scores, TIFF chose Mimsy due to its "more developed exhibitions module" as a way for Exhibitions to prioritise that

⁶² Melissa Neil, email message to author, May 20, 2014

department's needs. And since the FRL already had Star in place, the library's needs did not factor into the choice for Mimsy. ⁶³

When asked whether TIFF had considered open source systems or developing its own CMS in house, Neil stated it had not. One reason was the time commitment required of TIFF IT staff: "there would have been a fear that something like that could lead to a much longer process." Another reason was the familiarity with a known company name: "there was a real push to 'let's just use the system we've seen before,' use what other institutions are using: it's familiar, it's going to create a product that will look the way that other museums are expecting, the reports would look a certain way, and [TIFF senior staff] really equated this level of professionalism with using a known product."

The uncertain nature of the project funding was the last reason for not considering open source systems. Neil explained, "it was kind of a 'seize the moment' when you have an audience, someone's going to listen and maybe they will allow us to spend this money, so let's just do it.

Maybe you'll be halfway through the project and the money might disappear."

Neil saw the migration of data and metadata to Mimsy from the legacy system Star and from finding aids stored solely in various documents in a positive light. She attributed the relationship-building that Selago undertook with TIFF as the reason why it was successful. But, Neil also spoke of the challenges TIFF faced when it came to the data migration from Star:⁶⁴

...because Star wasn't relational, [Mimsy developers] referred to [Star] as having eight databases in it. The problem was there were film files, versus images, versus actual screening collections in Star. So you have all of these supposed databases and maybe

⁶⁴ Selago Design was the previous Mimsy vendor, prior to being bought out by Axiell.

⁶³ Melissa Neil, interview by author, Toronto, March 24, 2014.

there's [data for the same film] director in each one of them. All of that information then comes into Mimsy in our migration and we have six versions of the same [film] director in Mimsy: as a result, we are left with a lot of duplicate information. There was also a lot of talk and decision making about where data was going to go and how we were going to handle linking that data. There was the added complication of some data being catalogued to a library standard rather than to a museum standard, and getting Mimsy to deal with that.⁶⁵

To explain how such challenges as the above were met, Neil talked in detail about how staff from both departments were committed to a collaborative process where they pored over documents to find the best ways to resolve data conversion issues encountered through the migration. For Neil, the process of collaboratively streamlining the procedures of data entry, as well as resolving data clean-up issues was the best way to avoid any data issues in the future.

At the time of our interview in 2014, Neil described the migration process as finished. She added that, "for the most part it went very smoothly. And all the data is [in the system]. Everyone is finding everything. And [FRL staff] are still [cross-referencing the data with] Star to make sure that they are finding all the records."

Throughout the interview Neil emphasised how Mimsy would positively affect workflow. This is how she saw the situation:

It's going to have a huge impact on [Exhibitions and FRL staff's] ability to provide access, to let people search collections. But more importantly and less exciting to talk about- is that [Mimsy is] going to greatly increase our efficiency and the level of professionalism at which we can work thereby providing a much higher level of service to all of our clients. [At TIFF] there is a language [that] everyone is a client, even your colleagues, everyone is some sort of client or a customer.

Neil described efficiency, standardisation, and professionalism as the core issues, which all other procedures rested on.

⁶⁵ Melissa Neil, interview by author, Toronto, March 24, 2014.

What [Mimsy] is doing is really providing us the opportunity to standardise all of our procedures, forms, documentation, and all of our processes, so in that way, it is going to greatly improve our ability to care for a collection, our ability to keep proper records, and just generally force an improvement on workflow. It's undeniable to me that that's going to have a major impact...

Currently, TIFF is at the final stages of developing Mweb: the public face of Mimsy.

TIFF will publicly launch Mweb in fall 2015. FRL Manager and current acting Project Manager stated the following:

The first iteration of Mweb launched internally on April 29, 2015, in collaboration with "the Unseen Seen" exhibition opening reception. FRL staff provided a brief explanation on the goals of having an online searchable database, and provided ongoing demos throughout the event. The next phase is a roll-out to internal TIFF staff for feedback on user interface and search methods, followed by a third test phase with an external group of stakeholders. The external URL [will launch] in fall 2015.

Lovegrove Thomson also added that data cleanup and reconciliations are ongoing as a result of the data migrations from Star and special collection finding aids. She stated a three-year plan has been put into place for FRL staff to tackle any remaining data cleanup and reconciliations. Additionally, TIFF Mimsy manuals continue to be updated, mainly in response to changing work-flows.

Lovegrove Thomson provided the following figures to give an idea of the number of records now in Mimsy: 206,651 object records in the Library & Archives View; 364,088 people records; 67 acquisition records; 22 exhibition records in Exhibitions View; and 235 loan records.⁶⁷

_

⁶⁶ Michelle Lovegrove Thomson, email message to author, June 10, 2015.

⁶⁷ *Ibid*.

2. Ryerson University's Move to AtoM and Islandora

In February 2014, I interviewed Curtis Sassur, Coordinator of Ryerson University's Special Collections and Archives departments, and M.J. Suhonos, Digital Technologies Development Librarian at RULA to see how AtoM and Islandora functioned in practice. Sassur and Suhonos talked extensively about their experiences with Ryerson University's Archives and Special Collections' current CMS Minisis- and the reasons for moving to AtoM and Islandora—AtoM to meet the needs of special collections records, and Islandora for managing the library's digital content.

Like Star, Minisis is an older proprietary CMS developed in the 1970s. And as with other proprietary systems, Minisis has associated costs for installation and subsequent updates or customisation. AtoM and Islandora are on the other hand free and open source systems: access to both systems' source code for further use, modification, and redistribution is open to anyone.

Sassur stated his predecessor made the decision to migrate data from the institution's legacy software FoxPro to Minisis in 2008. The reasons for the move were given as the obsolescence of FoxPro at that time, and Sassur's predecessor's familiarity with Minisis from previous work at the Archives of Ontario.

When Sassur joined the Ryerson University's Archives and Special Collections team in 2011, he began to keep a list of issues the departments continued to encounter with Minisis.⁶⁸

This list included items such as frequent crashes, lack of interface customisation capability, and

_

⁶⁸ Curtis Sassur and M.J. Suhonos, interview by author, Toronto, February 28, 2014.

the lack of ability to close the system regularly. Sassur categorised the problems on two levels: "minor issues that are affecting individual users, and system-wide issues."

On a broader note, Sassur outlined the following regarding Minisis:

It's a rigid system. When you compare it to the possibility of an open-source system where you are basically able to customise it exactly as you want it seems very inflexible. Another issue: I had a really hard time getting my head around the nature of the relationship we have with the vendor. Every question that we had whether it was about functionality, or an issue we were having, or the need for training for the entry level staff that were using [Minisis], it was all framed in the proprietary relationship –it all has a dollar value. So that for me was problematic, on the ideological level.

These issues did not just affect Sassur and his staff at Ryerson University Archives and Special Collections, but also M.J. Suhonos' department. In order to attempt to resolve some of the recurring problems with Minisis, Archives and Special Collections would confer with Suhonos' department, part of the Library and Information Technology group (LIT) which provides technology support to RULA as a whole. Suhonos' understanding of the mounting frustrations at Archives and Special Collections was:

A lot of issues would happen, and before they would go to the vendor, they would come to us and say 'is this something that we can deal with on our own,' and quite frankly, some of them were increasing in frequency. And probably still increasing in frequency; it's not a good use of our time. That's how I looked at it. It was really a major inefficiency for us and continues to be, so it was a need to free up our work, our resources to focus on other things and to be more effective. In terms of identifying a candidate for replacement and the issues I saw with Minisis was what we call 'technical debt,' which is essentially that once you build a system, and are maintaining and modifying it over a long period of time you ultimately increasingly accrue this debt where it becomes too unwieldy to maintain [the system] any further, so the benefit of continuing to use it becomes outweighed by the maintenance of that debt.

The inefficiencies that Suhonos and Sassur encountered with Minisis ultimately led to the search for a new system to manage collections data. Sassur provided Suhonos with the list of issues that the Archives and Special Collections departments had encountered with Minisis in

2013. Sassur stated this was a signal that Ryerson University's Archives and Special Collections were ready for a change.

Sassur emailed me the list of issues he had gathered of issues encountered with Minisis:⁶⁹

- 1. Usability clunky interface, important fields hidden by multiple click-throughs, etc. different formatting of fields for Fonds/Series and File/Item level records is problematic.
- 2. System fails to shut down properly almost every time requiring a forced quit.
- 3. Unable to export enquiries data without crashing the system.
- 4. Cost for minor upgrades (e.g. restricted button) is prohibitive and out of proportion to the work required.
- 5. Process for deleting records is convoluted and often leads to 'ghost records' or deleted records returning to the database, or the need for expertise to be sought in order for [IT] to complete the delete task.
- 6. Possible issue around the number of concurrent licenses we have with upwards of 7 people now periodically working on our system at the same time.
- 7. Ideological issue of paying a vendor as opposed to seeking open-source in-house solutions.
- 8. Image hosting is complicated and involves having someone from LITs spend time on every single image that we wish to append to a record. This is needlessly complicated.
- 9. Limited capacity for hosting audio visual items appended to records.
- 10. Web Records display poorly.
- 11. Having to enter authority records in two places.
- 12. Time for requested upgrades and error fixes is very lengthy (we are still waiting for the ones we requested in august 2012).
- 13. Importing from excel sheets is overly complex and often results in errors.
- 14. Platform dependent and will not run on Mac OS or Linux OS.
- 15. No capacity for staff to input data into the system except on a workstation that has a licensed Minisis instance running. In contrast to AtoM, which allows web-remote data entry, this is a major weakness.

Sassur and Suhonos provided the following list of reasons why they opted for open source: reliability, lower associated costs, the service model attributed to the nature of F/OSS, customisation capabilities, efficiency, and compatibility with Ryerson University's core values of open access, continued research, collaboration, and experimentation. Sassur and Suhonos both

⁶⁹ Curtis Sassur, email message to author, June 8, 2015. Sassur referred to this list throughout the February 28, 2014 interview as the reasons why Ryerson University Archives and Special Collections began to look for an alternative CMS to that of Minisis. Sassur maintained in his June 8, 2015 email that the above list is indicative of his own views on the matter, and do not in any way reflect a broader critique of Minisis as a whole.

emphasised that the ability to share practice and code with other institutions in the implementation and development of an open-source CMS underscored "the value of the consortium—the Ontario College and University Library consortium." They cited this sharing of knowledge and practice among academic institutions as the reason they were both excited about the possibilities of open-source CMSs at RULA.

In discussion with Sassur, Suhonos subsequently chose two open-source systems, AtoM and Islandora. In terms of attributes, Suhonos chose AtoM for the archival and special collections data currently in Minisis because "AtoM is designed for standards-based archival description. [It is a] system for archival arrangement that supports all the archival standards with which we would want to comply. Most of our data in Minisis is predominantly [archival] descriptions and because it is mostly just descriptions and not a lot of [digital] content, [AtoM] is a good fit." On the other hand, Islandora was chosen for RULA's non-archival, non-special collections data because the system is "focused on managing digital content. [Islandora] is going to be for storing other kinds of digital content for the library: theses and dissertations are going into Islandora." Given Islandora's capabilities in accommodating rich digital content, Suhonos gave the example of videos which could be stored in the system later on.

Unlike proprietary systems, AtoM and Islandora are freely available for institutions to use, although if any added functionalities are required from AtoM for example, the system developer Artefactual can provide them at a cost, or RULA can develop them in house, utilising its own expertise in that area. The same applies to any technical support required. Sassur and Suhonos both stated that they valued the "flexibility to change course of action without having to wait for an external response, or support to happen."

Sassur and Suhonos also agreed that developing new functionalities or maintaining the system in-house was a positive thing for RULA. Sassur felt "you get added value, you don't just get a functionality that your staff puts in. And, the staff [have] that learning process of doing it, so you have a more enhanced staff as a result. There are some value-adds that are impossible to put a dollar figure on." Suhonos understood the issue as "I think of it as redirecting money that is currently being taken out of the institution and redirecting that into our own staff. I see it as reinforcement."

Migration of content was a big part of the move to AtoM and Islandora. Sassur described the migration process from Minisis to AtoM for Ryerson University Archives and Special Collections in similar terms as the migration undertaken by TIFF:

Basically we have a database system. We have our records and the information. And then we have the meta-information —that is the relationship of the records to each other and to the system. So when you're talking about getting your content out of a system and putting it in another, you have to be able to export [both the data and the metadata]. Because you have to accurately export all the data that's in all the individual records, and the relationships between the records to one another.

As for challenges in the migration of data and metadata from a proprietary system to an open-source system, Suhonos stated that it was as a result of Minisis using a proprietary data format in designing their systems. To make matters more difficult, migration of data and metadata from Minisis meant tackling issues that arose from "a botched transfer from FoxPro into Minisis. And that's not Minisis' fault. The data that we're trying to get out is already compromised." Sassur provided an example: "one of the limitations of FoxPro was that you could only have a title of about thirty characters, but in archival description the standard is that

⁷⁰ Curtis Sassur and M.J. Suhonos, interview by author, Toronto, February 28, 2014.

the title supplied is the title of the item. Some books' subtitles are seventy-character titles so they are just truncated [in FoxPro]. And so they got truncated into Minisis. We're dealing with titles that have been truncated by a legacy system."

Recent emails from Sassur indicate that "Ryerson's institutional repository of theses is now running in Islandora." Additionally, Sassur stated that the transfer of data from Minisis to AtoM was conducted entirely in-house by LITs and Special Collections staff and that Ryerson University's iteration of AtoM is also now live and managing Archives and Special Collections data. RULA has not requested any new functionalities from the system developer Artefactual for AtoM, nor have any new ones been developed in-house. ⁷¹

_

⁷¹ Curtis Sassur, email messages to author, June 8-August 26, 2015.

VII. Conclusion

Cultural heritage institutions are responding to changing technologies by transforming the ways they approach their collections and how they organise and manage their institutional data using digital applications. Mimsy, AtoM and Islandora are relational systems that can be customised to facilitate access to collections data, and increase efficiency as well as standardisation at a cultural heritage institution. Each system is better suited to one type of a cultural heritage institution (museum, archive, and library) than others: Mimsy is designed for museums, AtoM is specifically for archives, and Islandora for institutional repositories at libraries to manage digital assets. Each of these systems is better suited to one type of institution because the underlying data model and system interface relates to the specific cultural heritage context—the entities that represent the main tables in a data model created during the data modelling stage of database design are indicative of the institution's information management requirements.

In addition to the data models and cultural heritage contexts, each institution analysed the proprietary and open-source options and selected the one that fit their requirements. The selection process and the choices they made can be used to better understand the CMS landscape and the pros and cons of each type of option - proprietary or open-source. TIFF and Ryerson University Archives and Special Collections provided differing reasons why they chose proprietary (TIFF - Mimsy) or open-source systems (RULA – AtoM and Islandora) for managing their collections:

At TIFF, the reasons for not choosing an open source system are tied to precarious project funding, lack of available IT staffing resources for the project, and the expectation to go with a known company name that has provided products and services to other respected cultural heritage institutions. Thus for TIFF a turn to open source was seen as a risk of short-term project funding in light of better known proprietary options available.

For Ryerson University's Archives and Special Collections, values of open access, research, and collaboration conflicted with that of the previous proprietary system in place which did not allow for easy access to the system, let alone the content within it. Furthermore, the inefficiencies of Minisis as well as access to the IT staffing resources necessary for the development and implementation of a new CMS made the decision to explore open-source options a natural one.

Although each institution had its own set of reasons for choosing a new CMS, both wanted similar results. Neil's understanding was that TIFF was not only looking to increase further access to its collections for staff and the public, but that Mimsy would also increase TIFF's level of standardisation, professionalism, and efficiency. These expected improvements at TIFF would strengthen the institution's place among cultural heritage counterparts.

At Ryerson University, Sassur and Suhonos' hopes for the implementation of opensource systems AtoM and Islandora were not altogether different: that RULA would ultimately have increased access to its collections, a more efficient workflow, and better adherence to industry standards. Additionally, Sassur and Suhonos hoped that with the development and implementation of AtoM in particular, they would take part in collaborations with other institutions at the forefront of research in the library sector. For RULA, this would also mean better support from the consortium of which the school is a part. Overall, Sassur and Suhonos saw the turn to open source in general, and Islandora and AtoM in particular, as an investment in the institution itself as well as its place amongst other leading universities of its kind.

These findings on proprietary and open source systems indicate that both models are capable of meeting the basic data and collections management needs of cultural heritage institutions –museums, libraries, and archives. Although open-source systems for collections management are more recent additions to the well-known lists of proprietary off-the-shelf product names, open-source systems such as AtoM and Islandora are establishing themselves as collections management and institutional repository contenders no less worthy of attention and use. Additionally, the long-term development and community support that is enabled by the open-source model provides incentives for taking a closer look at these options. The increasing availability of flexible technical support options for smaller-sized institutions further advances the case of open-source solutions.

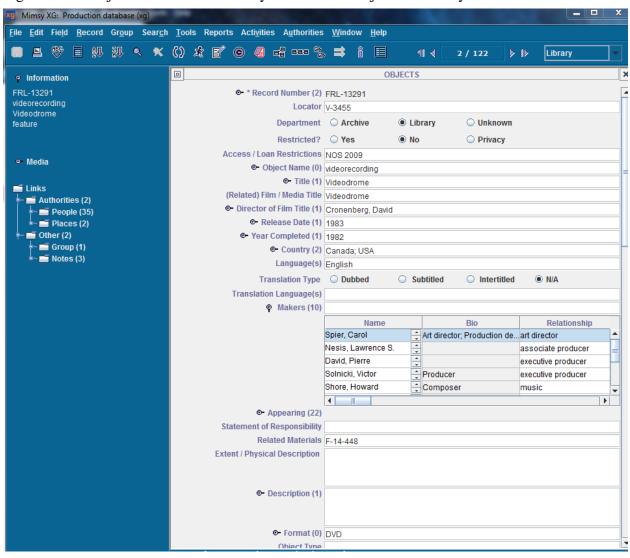
To conclude, the effectiveness of a CMS is ultimately tied to the type of cultural heritage institution where it is implemented as well as the data model that system is based on to meet institutional needs and requirements. Institutional needs and requirements differ from one cultural heritage institution to the next due to differing mandates, varied resources available, and diverse collections requiring enhanced approaches to access, preservation, and digital management. Proprietary and open-source systems answer these data and cultural heritage

requirements in different economic and resource development models. The case studies presented in this thesis provide an overview of these issues and highlight the complicated selection and decision making process which determines the CMS solution that addresses cultural heritage information requirements in a digital era. In the process the CMS landscape is explored, highlighting diverse contexts (library, archive, museums) and models (data, social, and economic), that can be used to assist institutions in developing their own road-map to implementing CMS solutions.

Appendices

Screenshots of Mimsy at TIFF

Fig. 1.1 – An object record in the Library View of the Objects Authority



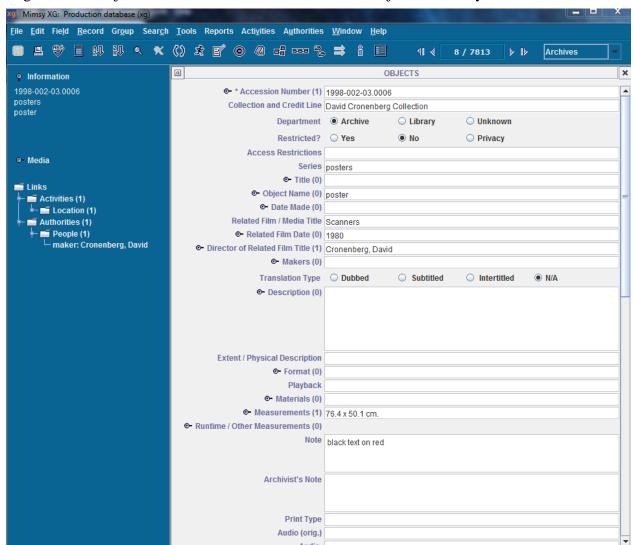


Fig. 1.2 – An object record in the Archives View of the Objects Authority

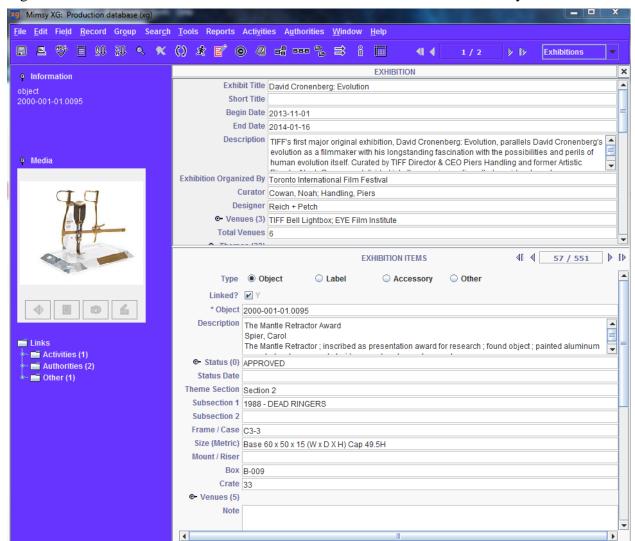


Fig. 1.3 – An exhibition record in Exhibitions View of the Exhibitions Authority

Screenshots of AtoM at Ryerson University's Archives and Special Collections

Fig. 2.1 – Object record entry page with contracted drop-down fields

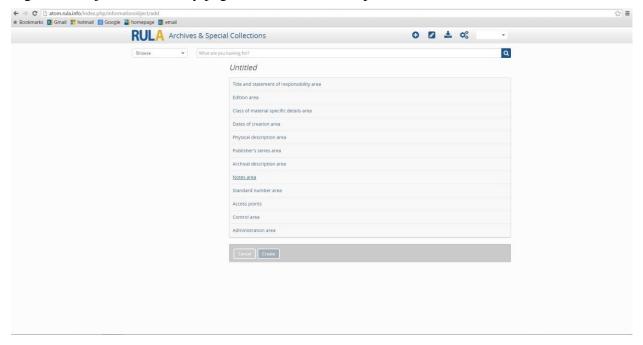


Fig. 2.2 – Object record entry page with expanded drop-down fields

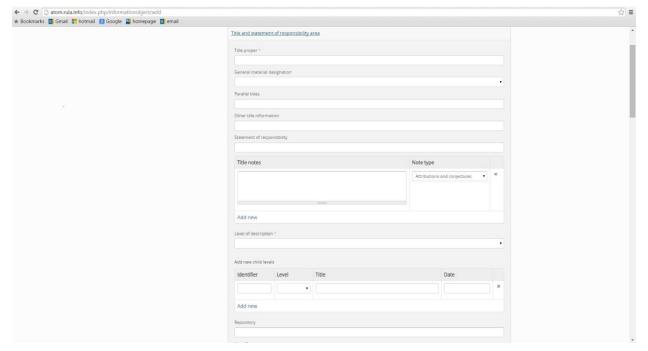


Fig. 2.3 – Object record entry page cont'd.

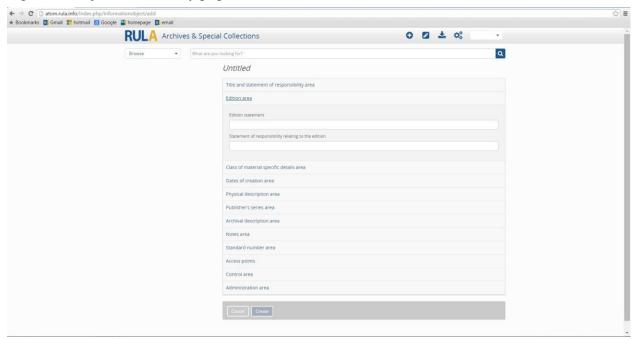


Fig. 2.4 – Object record entry page cont'd.

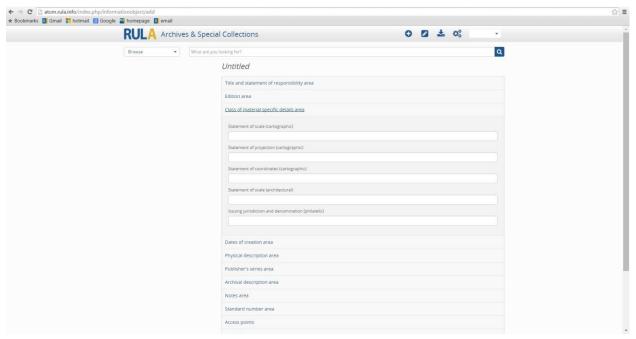


Fig. 2.5 – Object record entry page cont'd.

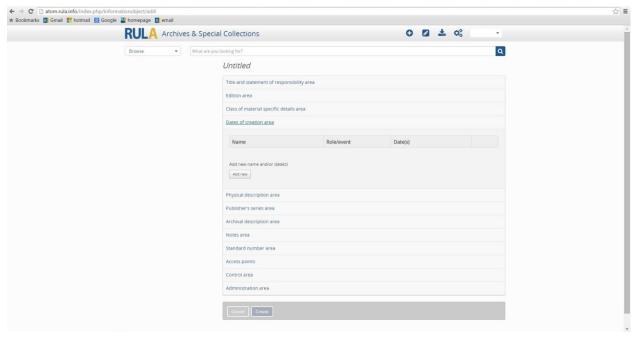


Fig. 2.6 – Object record entry page cont'd.

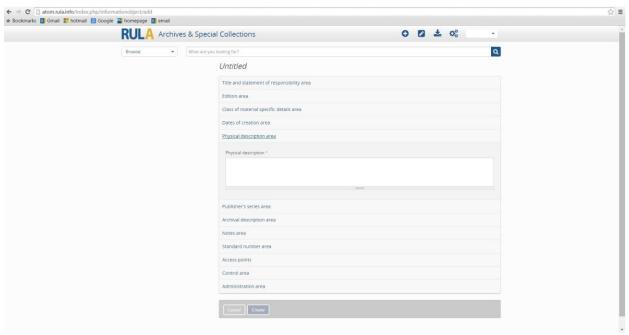


Fig. 2.7 – Object record entry page cont'd.

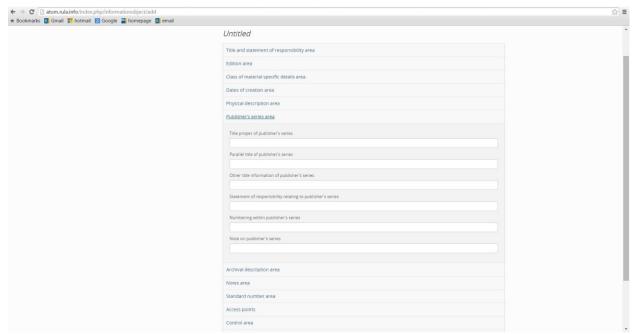


Fig. 2.8 – Object record entry page cont'd.

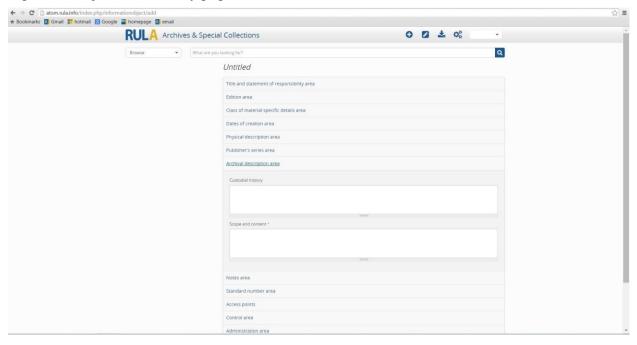


Fig. 2.9 – Object record entry page cont'd.

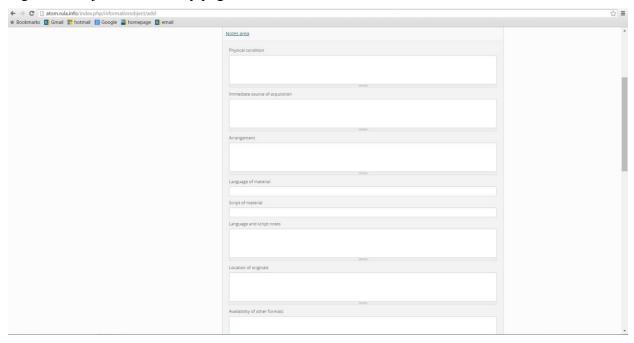


Fig. 2.10 – Object record entry page cont'd.

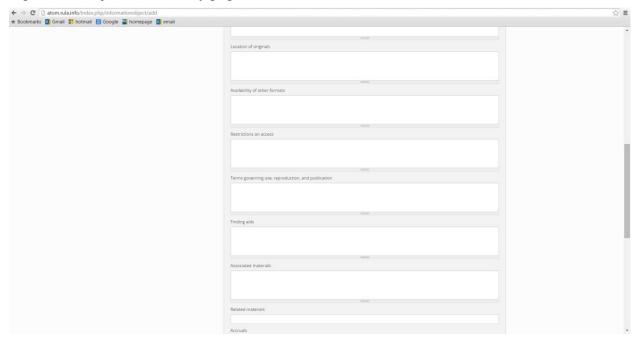


Fig. 2.11 – Object record entry page cont'd.

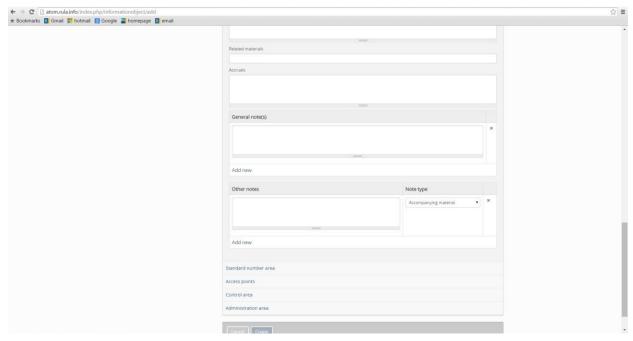


Fig. 2.12 – Object record entry page cont'd.

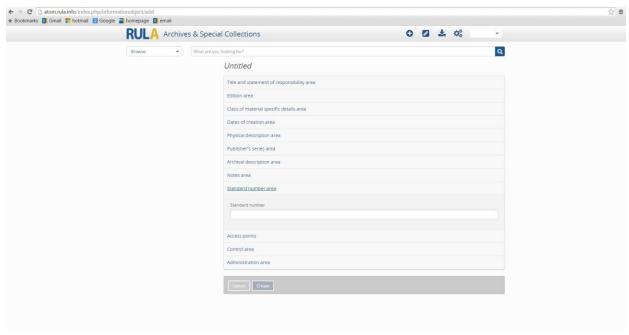


Fig. 2.13 – Object record entry page cont'd.

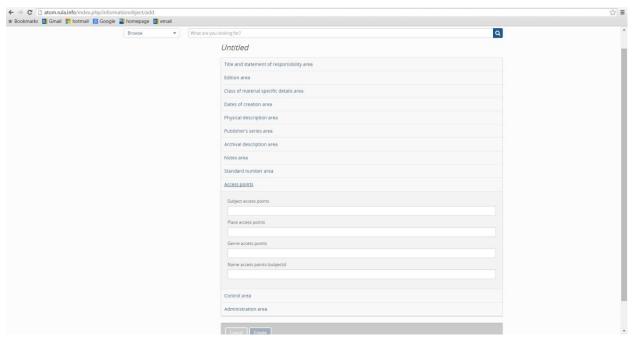


Fig. 2.14 – Object record entry page cont'd.

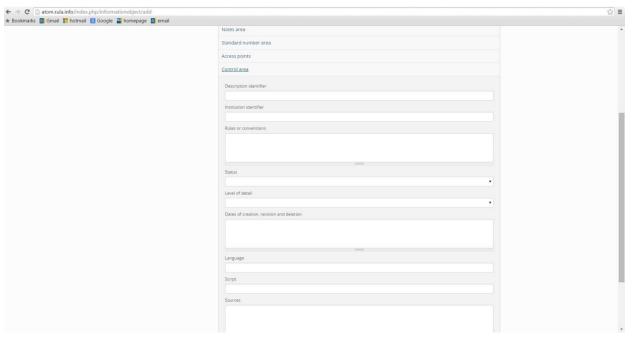


Fig. 2.15 – Object record entry page cont'd.

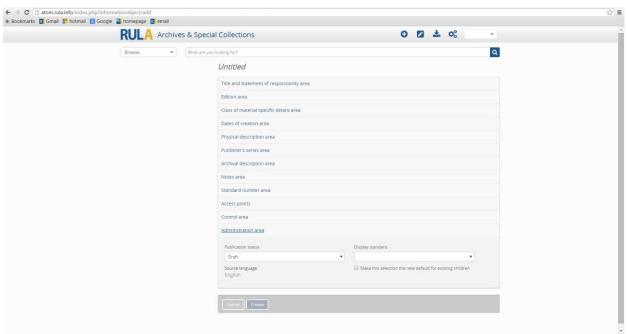
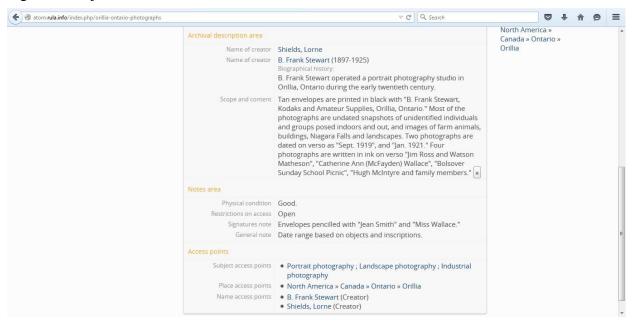


Fig. 3.1 – An object record in AtoM



Fig. 3.2 – Object record cont'd.



Screenshots of Islandora at RULA

Fig. 4.1 – Object record entry page for Theses and Dissertations with expanded drop-down fields

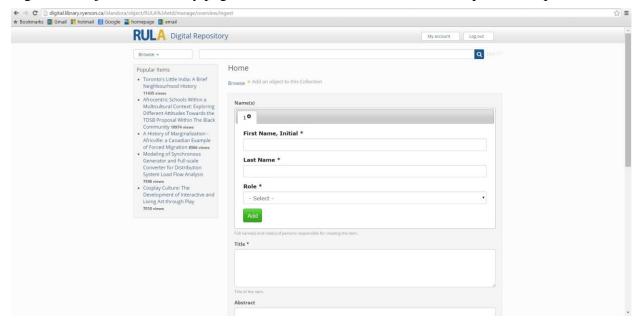


Fig. 4.2 – Object record entry page cont'd.

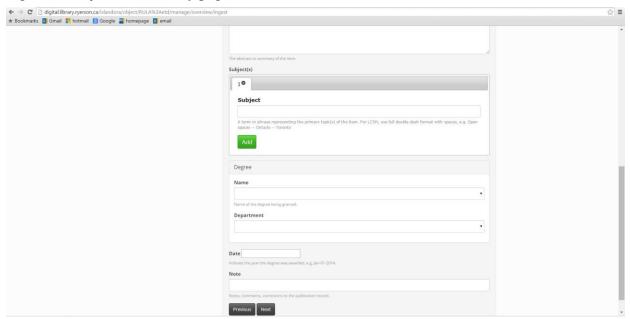


Fig. 5.1 – An object record in Islandora

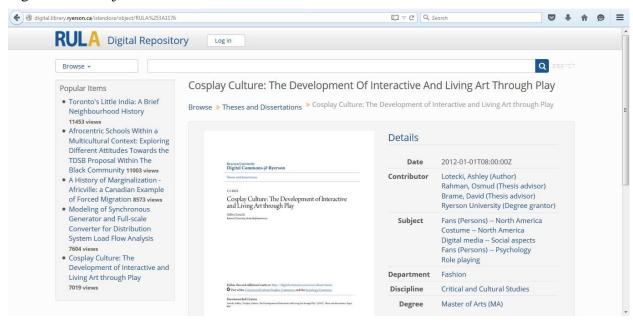
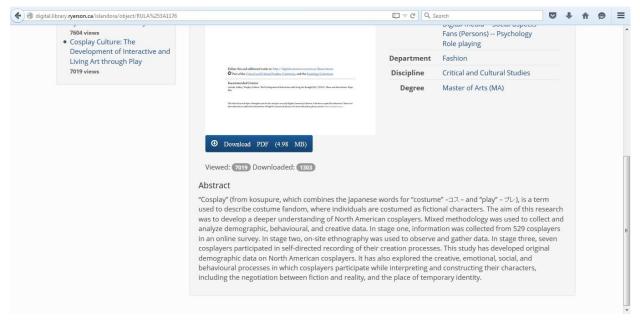


Fig. 5.2 – Object record cont'd.



Bibliography

Resources on Databases

- Carpinone, Elana C. *Museum Collections Management Systems: One Size does NOT Fit All.* MA thesis, Seton Hall University, 2010.
- Date, C.J. The Database Relational Model. USA: Addison Wesley Longman, Inc., 2001.
- Levene, Mark and George Loizou. *A Guided Tour of Relational Databases and Beyond*. London: Springer-Verlag, 1999.
- Paradaens, Jan, Paul De Bra. Marc Gyssens, and Dirk Van Gucht, Monographs on Theoretical Computer Science, vol. 17, *The Structure of the Relational Database Model*. Berlin: Springer-Verlag, 1989.
- Perry, Shannon Leigh. So, You Want a Database?: A Beginner's Guide to Database Technology for Small Cultural Institutions. MA thesis, Ryerson University, 2006.

Cultural Heritage Institution Mandates and FAQs

- Archives and Special Collections. *Archives*. Accessed May 15, 2015. http://library.ryerson.ca/asc/archives/.
- Archives and Special Collections. *A&SC Collections Policy information for Potential Donors*. Accessed May 15, 2015. http://library.ryerson.ca/asc/donations/.
- TIFF. Film Reference Library FAQ. Accessed November 24, 2013, http://tiff.net/filmreferencelibrary/library/faq.
- TIFF. Our Story. Accessed May 15, 2015, http://tiff.net/explore/about.
- Special Collections, *Archives and Special Collections*, Accessed May 15, 2015. http://library.ryerson.ca/asc/sc/.

Resources for Cultural Heritage Institutions

- Canadian Heritage Information Network. *Collections Management Software Criteria Checklist* 2012. Accessed June 14, 2015, http://www.rcip-chin.gc.ca/carrefour-du-savoir-knowledge-exchange/criteres gestion collections management checklist-eng.jsp.
- Collections Trust. *Choose a CMS*. Accessed June 14, 2015, http://www.collectionstrust.org.uk/choose-a-cms.

- Gill, Tony, Anne J. Gilliland, Maureen Whalen, and Mary S. Woodley, *Introduction to Metadata*, edited by Murtha Baca. Los Angeles: Getty Publications, 2008. http://www.getty.edu/research/publications/electronic_publications/intrometadata/index.html.
- Note, Margot. *Managing Image Collections: A Practical Guide*. Oxford: Chandos Publishing, 2011.
- Poma Swank, Annamaria. *Collection Management Systems Report*. Report presented at the Fondazione Rinasimento Digitale, Italy, 2008.
- van Hooland, Seth and Ruben Verborgh. *Linked Data for Libraries, Archives, and Museums: How to Clean, Link and Publish Your Metadata*. Chicago: Neal-Schuman, 2014.

Resources on Proprietary and Open-Source Software

- AtoM. *Community: Get Support*. Accessed August 28, 2015. https://www.accesstomemory.org/en/community/support/.
- AtoM. *Entity Types*. Accessed August 28, 2015, https://www.accesstomemory.org/en/docs/2.3/user-manual/overview/entity-types/#entity-types.
- AtoM. What is AtoM?. Accessed July 14, 2015. https://www.accesstomemory.org/en/docs/2.2/user-manual/overview/intro/.
- CHIN Professional Exchange. *An Introduction to Free and Open Source Software for Museums*. Accessed February 1, 2014. http://www.pro.rcip-chin.gc.ca/carrefour-du-savoir-knowledge-exchange/intro_code_source_ouvert-intro_open_source-eng.jsp.
- DiBona, Chris, Mark Stone and Danese Cooper, eds., *Open Sources 2.0: The Continuing Evolution*. Cambridge: O'Reilly Media, 2005.
- Discoverygarden: Managing Digital Content. Why are Organisations Migrating to Islandora from Other Digital Repository Systems?. Accessed July 14, 2015. http://www.discoverygarden.ca/islandora-overview/.
- Feller, Joseph, Brian Fitzgerald, Scott A. Hissam, and Karim R. Lakhani, eds., *Perspectives on Free and Open Source Software*. Cambridge: MIT Press, 2005.
- ICA-AtoM. *Open Source Archival Description Software*. Accessed February 1, 2014. https://www.ica-atom.org/download/ICA-AtoM_JBushey.pdf.
- Islandora. *Islandora Community*. Accessed August 28, 2015. https://wiki.duraspace.org/display/ISLANDORA715/Islandora+Community.
- MIMSY. *Mimsy XG*. Accessed July 14, 2015. http://www.willo.com/mimsy_xg/.

MINISIS. Minisis Inc. Accessed July 14, 2015. http://www.minisisinc.com/index.html.

Selago Design, Inc. *Portfolio: Mimsy XG*. Accessed July 14, 2015. www.selagodesign.com/portfolio/mimsyxg/index.php.

Weber, Steven. The Success of Open Source. Cambridge, MA: Harvard University Press, 2004.