A Descriptive Survey of Illinois’ Digital Collections

by Miriam Lytle, Marie Martino and Justine Wagner-Mackow

Dominican University
Graduate School of Information and Library Science

Spring 2009

Introduction

A Descriptive Survey of Illinois’ Digital Collections research project endeavors to understand digital collections as they exist in the State of Illinois in spring 2009. Public, academic, special and museum libraries completed a questionnaire regarding their organizational type, collection description, selection process and planning, searchability and sharability, marketing, maintenance and sustainability. The result is a benchmark, an actual snapshot in time, because digital collections and digital libraries are rapidly evolving in the field of library and information science. A follow-up survey in subsequent years would be appropriate for tracking how digital collections develop with technological advances, user interface, funding and time-tested experience.

Background and Literature Review

Digital library research and development began in 1994 with $24.4 million in federal funding for a joint initiative of the National Science Foundation, Department of Defense Advanced Research Projects Agency and the National Aeronautics and Space Administration (NASA). Six universities (Carnegie Mellon, California-Berkeley, California-Santa Barbara, University of Illinois-Urbana Champaign, University of Michigan and Stanford) shaped computer science architecture and information retrieval systems that catapulted libraries into the 21st century (Besser 2002). Fifteen years later, we seek to define a digital collection for the purpose of this study.

Discovering the ambiguity, differences and interchangeability between the terms digital collections and digital libraries is a separate etymological study in itself. The entanglement of terms, jargon and inconsistencies exemplify growing pains in the changing library and information science field (Borgman 1999) which continues to afflict us through the design, data collection and findings of this study. A digital library is defined as “digitally born” archives to one organization (Stielow 2003), yet means using electronic databases to another (Arms 2000), and to a third, the interoperability and overlay of services (Digital Library Foundation 2006). William Arms, creator of D-Lib URL, defined a digital library as “a managed collection of information, with associated services, where information is stored in digital formats and accessible over a network.” While Arms stresses the importance of the word “managed,” a
digital library by his definition includes collections such as NetLibrary, My Media Mall and online databases (Arms 2000). In contrast, a list of attributes compiled by the Digital Library Federation’s (DLF) states that leasing digital resources does not make a library digital (Digital Library Foundation 2006).

Adopting the Digital Library Federation’s definition for the purpose of our research allows us to distinguish between a digital library and a digital collection.

Digital libraries:
- are organizations that provide resources that are readily and economically available
- provide services to a defined community or set of communities, and are thereby defined and measured in relation to the communities they serve
- have specialized staff
- select, structure, and interpret collections of digital works
- distribute and offer intellectual access to these resources
- preserve the integrity of the collections of digital works
- ensure that these resources will persist over time

A literature review reveals that the definition of the term “archive” bleeds heavily into the area of collections. The term digital archive can be “born digital,” such as digital photographs, e-newsletters and other types of electronically distributed texts and images (Stielow 2003). However, Preserving Digital Information: Report of the Task Force on the Archiving of Digital Information (Waters & Garrett 1996) regarded a digital archive by its archival nature and by preservation of the work collected.

**Conceptualization and Sampling**

Like Dorothy Gale of Kansas we eventually discovered that we possessed all that we needed. Rather than tapping our ruby red slippers together, however, we phoned the Illinois State Library in Springfield. Library Service and Technology Act (LSTA) grants consultant Karen Egan patiently listened to our proposed research, expressed enthusiasm and immediately forwarded the call to Alyce Scott, the State’s digital imaging program manager. Suddenly we were back in Kansas and on the road to defining digital collections for the purpose of our study. Through phone conversations and emails, Alyce Scott graciously assisted us in determining the parameters of digital collections and digital archives according to the State of Illinois. If not for their kindnesses, we might still be lost in Digital Oz.

The Illinois Digital Archives (IDA) serves the Illinois State Library’s mission to promote “excellence in information access and innovative services for government, libraries and people” shared via the Internet, especially local, state and national history materials in various formats from the State’s libraries, museums, archives and other cultural institutions (IDA Collection Development Policy 2007-2008). The LSTA Digital Imaging Grant (FY2010) provided us with the parameters for researching Illinois’ digital collections.

With a grasp of definitions, we chose existing digital collections in Illinois for our sample frame. We also ascertained that digital archives and collections are “raw” with limited and varying degrees of interoperability (Besser 2002), which differed from the Digital Library Federation’s definition of digital libraries. The difference is summed up by Clifford Lynch’s *First Monday*
observations in "Digital Collections, Digital Libraries and the Digitization of Cultural Heritage Information" (Lynch 2002):

“One of the things that I think is starting to become clear is that digital collections and digital libraries aren't the same thing. This is a crucial observation...Digital libraries seem to me in some sense to be enormously more complex, enormously more open-ended. ...Digital libraries are as rich as our visions about how we can use and reuse digital information. They're as rich as the conversations we can imagine between books. They're as powerful as the linkages we can imagine creating as we amass material and in the digital world it becomes greater than the sum of its parts. And perhaps it's good that we're not ready to produce a best practices guide for true digital libraries (as opposed to digital collections masquerading under the label of digital libraries), that everything is still very much open to trying new things and exploring new ideas in this area...(and) how to continue to evolve our thinking about digital collections and our practice in creating and maintaining them.”

On February 12, 2009, we also started a team project blog to help us more efficiently share ideas, data, timeframes and task lists. We shared our blog with a handful of colleagues who generously offered their support as the project moved through conceptualization to the final report.

During our literature review, we learned that while there had been usability studies, there were no descriptive surveys known to date on Illinois’ digital collections (Scott 2009). We searched exhaustively to identify and count Illinois’ digital collections to determine the population. More than 292 digital collections/archives were evident as of February 28, 2009. We say “more than” because some collections are “faculty publications” and “university archives” while others are on password-protected websites, inaccessible to lurkers the likes of us. We counted nine host sites and their collections, as follows: Barrington Area PL’s collections (n=3); CARLI - Consortium of Academic and Research Libraries in Illinois (n=43); DePaul Universities (n=68 +); Illinois State Library Illinois Digital Archives (n=53); Illinois State University History, Milner Library (n=5); Lincoln Trail LS (n=2); NSLS - North Suburban Library System Digital Past (n=85); University Library, University of Illinois at Urbana-Champaign (n=29); Upper Mississippi River Valley, not including Iowa (n=4).

However, with new collections processed daily, we also discovered that setting the sample frame was like trying to count a colony of busy ants. We abandoned the tally, deciding it provided better objectivity and authenticity for organizations to report directly during the data collection phase of our research. This perception was confirmed when we read the March 1st North Suburban Library System’s Digital Past blog stating, “We’ve seen a lot of growth lately…building up (not so slowly) to the 100,000 item mark in Digital Past! As of today, we’re just over 92,300.” A day later Kay Schlumpf, Digital Past Coordinator at NSLS posted on our project blog, “Just so you know, we’re well past the 100,000 mark now. Right now, this morning, we’re at 135,865 with 40 institutions.”

Also, we were interested in learning how digital collections were selected, constructed and maintained, as well as their content, significance and formats. We were curious about the technical aspects of managing a digital collection, how one is able to locate a specific collection on the Internet and how they are marketed to the public. Considering digital collections depend on careful life cycle management (Smith 2001), obviously, this weighty topic spawned many questions. We not only sought out information about best practices and established standards that
foster quality, consistency, and interoperability across multiple systems, but also an overarching model in which to consider digital collections. Having such an overview would help us identify the variables and aid us in the design and organization of a simple random probability survey with Illinois organizations and their collections as the units of analysis.

**Methodology, Instrument Design and Operationalization**

Our objective was to conduct reliable research using an appropriate set of questions to sample the State’s collections in the cyberscape. The National Information Standards Organization’s (NISO) publication “A Framework of Guidance for Building Good Digital Collections” provided us with a structure in which to examine the implementation of digital collections, as well as a set of criteria that promotes the quality, consistency, interoperability, accessibility, and sustainability of digital collections over time. NISO's list of guiding principles and initiatives centered on four interrelated core areas: collection programs and projects, digital objects, collections, and metadata (NISO 2007, p. 7). Focusing on key points from each sphere, we developed a fairly comprehensive survey instrument to determine digitizing projects by Illinois libraries and other cultural institutions by 2009 standards.

In order to gather qualitative and quantitative data, we designed an online questionnaire as our research instrument. An added benefit was that the individual respondents remained anonymous, therefore eliminating the need for Institutional Review Board consent forms. Before building our questionnaire, we tested free versions of Google Docs, Wufoo, and SurveyMonkey. Although each of these programs has advantages, we chose SurveyMonkey, which, after six test versions, proved that researchers and respondents alike could easily navigate the intuitively designed and logically arranged questionnaire with contingencies. Additionally, the program collected and analyzed the data with graphs. For the purpose of this study, we upgraded to SurveyMonkey’s professional edition.

Many of the specific elements in the questionnaire overlap with the four NISO-identified areas, so we organized for intuitiveness, fluidity, validity and transferability by dividing it into four sections: Collection Description, Standards and Sustainability, Findability and Sharability, and Usage and Marketing. Most of the questions were closed-ended with finite answers with only the first three questions required, while the remaining twenty-two questions were optional. Where appropriate, we provided spaces for answers not listed by providing an “Other (please specify)” write-in option. Anticipating a wide range of digital expertise levels among our respondents, and being that a significant portion of the questions were technical, we decided to add the “I don’t know” response option to appropriate questions. The final open-ended question functioned as a space for respondents to substantially communicate anything more about their organization’s digital collection(s) that we had not already asked (Appendix C).
Collection Description
The Digital Collection Description section of our survey provided us with a baseline description of digital collections and their content. The first couple of questions asked the institution’s name and type. The following five questions focused on the NISO areas of digital objects and collections. Respondents were asked to identify the number of collections that their organization maintains in an open-ended question to accommodate those who may have collections in various states of being. The subsequent four questions dealing with collection scope and subject matter, and digital object format and material types, were based on Illinois State Library LSTA Digital Imaging Grant (FY2010) criteria on general subject headings and format. We then asked respondents if they host their collection(s) on their own organization server; if they do not, we asked respondents to identify the organization that hosts their collection, allowing for alternate hosting sites by providing an “Other: (Please specify)” write-in option. The nine hosting institutions that we provided by name on the questionnaire were chosen because they were identified on the Digital Illinois website as hosting institutions (Digital Illinois, 2007).

Standards & Sustainability and Findability & Sharability
The remaining fifteen questions (excluding the last), spread out across the three sections of our questionnaire, addressed the NISO-identified areas of collection programs and projects and metadata. Collection programs and projects allude to digital library initiatives and activities that span the whole lifecycle of the digital collection. This includes everything from planning and policy development, to active data management and preservation, to services, marketing, and evaluation (NISO 2007, p. 91). “Metadata,” within the context of this survey, refers to descriptive metadata, or data about digital objects or collections that allows for resource identification, retrieval, and discovery (NISO 2004, p. 3). Our interest lies in how it relates to digital collection standards, findability and sharability.

Within the Standards and Sustainability and Findability and Sharability sections, we asked questions related to digital library standards and best practices, which are inextricably linked to the overall impact and sustainability of digital library initiatives. As Washington State Library’s Digital Best Practices webpage states, “... the opportunity exists to discover new ways to accomplish the tasks and share proven methods. It is to everyone's advantage to maintain a clearinghouse of best practices to share with project participants and to disseminate to other organizations” (Washington State Library). The remaining questions in these sections centered on different components of the digitization process and the creation and use of metadata as it relates, ultimately, to the ways users can access and interact with digital collections. Such standards are key in the digital environment. “…[T]he function is to try to unify the representation, manipulation or transmission of some piece of information so that two or more different systems can ‘understand’ it the same way. [Standards] are the basis of interoperability, portability, modularity, building blocks, objects, and all the other names invented to describe how two pieces of software should be able to simply work together” (Noerr 2003, p. 97). We provided respondents with a short list composed of several, well-known institution-based standards. Recognizing that the list was not definitive, we also offered an “Other: (Please specify)” write-in option.

The employment of optical character recognition (OCR) software makes it possible for users to search the actual text that a digital object may contain within the digital collection database, thus increasing their access to materials. We included a comprehensive list of known metadata
standards and inquired about different aspects of metadata – standards, Open Archive Initiative (OAI) protocol for metadata harvesting, and including metadata tags in xhtml/xml Web documents. Understanding that institutions may modify, use a combination of standards or use an internal system, we asked respondents to check all that apply and presented them with an “Other: (Please specify)” write-in option. The OAI protocol functions as a “tool to move and aggregate metadata from diverse institutions.” Via the OAI protocol, organizations can expose their metadata so that an aggregator service can incorporate the organization’s metadata into a database, making it accessible to users. Some examples of such aggregators include: the National Science Digital Library (NSDL), OAIster, American South.org, and the IMLS Digital Content Gateway (DLF 2007). Finally, the use of metadata tags in xhtml Web documents connected to database portals or even organizational homepages makes it possible for users to find digital collections by way of common search engines, like Google or Yahoo. Again, because of the technical nature of these questions, we found it especially useful to include “I don’t know” as an alternative to the standard “Yes” or “No” options.

Usage and Marketing
Lastly, in the survey’s final section, we address the issues of user services, marketing and collaboration efforts built around digital collections. “Marketing should not be an add-on, but an essential part of building good digital collections, and funds for anticipated marketing expenses should be included in project and program budgets” (NISO 2007, p. 98). A critical facet, we wanted to explore how libraries and cultural institutions integrate related activities into their services. Are they promoting their collections on the Web through websites and/or through Web 2.0 applications like Wikipedia, Flickr, or Facebook? Are they publicizing their digital collections through traditional means such as print and programming? Respondents were allowed to check all the options, as well as use the “Other: (Please specify)” write-in option. Are organizations engaging users with participatory or social tools like tagging through the user interface? Also, are organizations working collaboratively? How are they harnessing the power of these relationships to boost their digital collections? Are organizations evaluating collection usage? Finally, what preservation efforts are in place? Are Illinois institutions backing-up these large amounts of data with external services, such as LOCKSS or Portico, or are they taking this effort up themselves and what method are they utilizing?

Through a series of rewrites, the help of digital professionals and sheer determination, we narrowed our questionnaire to 25 questions. Our test volunteers reported it to us that the questionnaire took between five and ten minutes to complete.

Data Collection, Analysis and Findings

The survey was promoted by the Illinois State Library e-newsletter, Dr. Michael Stephens’ Tame the Web blog, WebJunction Illinois announcements and blog, and the following listervs: CARLI Collections and Content Group, CARLI Digital Collections, CONTENTdm, Digital Library Federation, Library and Information Technology Association, Metropolitan Library System, North Suburban Library System.

The survey was available online for three weeks and elicited 54 responses. Initial collection analysis in SurveyMonkey generated response percentages, response count, answered/skipped
questions, and graphs. Each respondent’s open-ended comments were printed out for open coding and were evaluated for inter-reliability by our team. While 54 institutions responded to our questionnaire during the data collection period, eleven respondents were eliminated from the final analysis for various reasons to ensure analysis reliability, dependability, validity and authenticity. Three out-of-state organizations were outside the study sample, several respondents had “0” collections, one respondent answered based on their NetLibrary, and others had significant missing data. Three organizations chose to remain anonymous, but their questionnaires provided reliable appropriate data and were therefore included. Of the 44 reliable questionnaires analyzed, 34 (77.3%) were completely finished. After removing eleven of the 54 surveys, we ran the data analysis for reliability and validity again.

**Key Findings: Digital Collection Description**
The following statistics and charts provide the key findings of this research project. Among the 44 reliable questionnaire respondents, a significantly higher percentage of academic libraries and universities responded to the survey, applicable to statewide digital collection generalizability (chart below).

**Question: What is your organization type?**

![Bar Chart](chart_url)
Collections that are selected to be digitized that have the highest significance to their immediate community are priorities (chart below). Most respondents had six or fewer digitized collections (n=33) but ranged up to 30 collections.

**Question: What is your organization’s collection(s) significance? (Check all that apply.)**

![Bar graph showing significance by percentage: Local is 84.1%, National is 43.2%, State is 52.3%, International is 36.4%.]

General subject areas ranked local history first, followed by newspapers and arts/architecture. Photographs/photographic negatives at 71.5% were the most popular types of materials digitized, followed by books at 54.5%. Format types reinforced these results with 79.5% image, and 77.3% text, 34.1% audio and 20.5% video (chart below).

**Question: Your organization’s digital collection(s) is comprised of the following formats (please choose all that apply):**

![Line graph showing format distribution: Text is 77.3%, Image is 79.5%, Audio is 34.1%, Video is 20.5%, Other is 2.3%.]
Outside institutions or agencies host 67.4% of respondents’ digital collections, with CARLI at 45.2% and Illinois Digital Archives at 32.3% (chart below).

**Question:** Who hosts your organization’s digital collection(s)?

**Key Findings: Selection Process and Planning**
Only 16.7% of respondents are aware of having a digital collection development policy (chart below).

**Question:** Does your organization have an official digital collection development policy?
Key Findings: Selection Process and Planning, continued
The Illinois State Library’s best practices are adhered to by 48.8% of survey respondents, followed by the Library of Congress best practices at 26.8%. Funding for digitizing is 73% internal budget and 53.7% LSTA grants, although there is budget overlap in some instances (chart below).

Question: How is your organization’s digital collection(s) funded? Please check all that apply.

Key Findings: Searchability and Sharing / User Access
Optical Character Recognition is implemented by 43.2% of the respondents (chart below).

Question: Do you employ optical character recognition (OCR) to make text documents searchable?
Key Findings: Searchability and Sharing / User Access, continued
Dublin Core is the metadata standard for 64.9% of respondents (chart below). Compliance with Open Archives Initiative (OAI) protocol to allow for metadata harvesting is 37.8% while 24.3% use metadata tags for website findability, according to respondents who answered.

**Question: What metadata standard does your organization utilize?**

<table>
<thead>
<tr>
<th>Metadata Standard</th>
<th>Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dublin Core</td>
<td>64.9</td>
</tr>
<tr>
<td>Encoded Archival Description (EAD)</td>
<td>16.2</td>
</tr>
<tr>
<td>Machine-Readable Cataloging Record (MARC)</td>
<td>29.7</td>
</tr>
<tr>
<td>Visual Resources Association (VRA Core 3)</td>
<td>8.1</td>
</tr>
<tr>
<td>Other</td>
<td>13.5</td>
</tr>
</tbody>
</table>

Key Findings: Marketing the Digital Collection
The marketing and collaborative questions were more in-depth. Most organizations work collaboratively with other libraries (59.4%) for collection procurement (42.3%) and over a third work together on funding.

Digital collections are marketed primarily by the organization’s website at 88%, followed by print. However, Web 2.0 social networking sites are being used by some (chart on next page).
Question: If your organization does market their digital collection(s), how do they do it? Please check all that apply.

Key Findings: Maintenance and Sustainability
More than half of the respondents chose external storage for backing up their collections (chart on next page).

It is evident that more sustainability planning needs to be implemented based on only 28.6% of respondents know of their organizations’ sustainability plans, and 17.1% have allocated budgets. (chart on next page).

Usage statistics on their digital collections are kept by 56.8% of the respondents.
Question: How do you back up your collection? Please check all that apply.

- **Other**: 13.5%
- **Tape backup on server**: 40.5%
- **Portico**: 2.7%
- **OCLC Digital Archive Service**: 0%
- **LOCKSS (Lots of Copies Keeps Stuff Safe)**: 10.8%
- **Internal computer storage**: 27%
- **External Hard Drive**: 29.7%
- **External Storage, such as CD on site or off site**: 51.4%

Question: What does your organization have in place for the sustainability of the digital collection(s)? Please check all that apply.

- **Plan**: 28.6%
- **Budget**: 17.1%
- **I don't know**: 54.3%
- **Other**: 11.4%
Key Comments
Nine respondents answered the last open-ended question, providing qualitative data. Even from so small a sample, we identified trends of growth, change and technological advancement. Several respondents were eager to tell us the stage of their digital program, especially those just beginning or in transitional stages, “We are just beginning,” or “We are currently attempting to shift from a project based environment with ad hoc procedures for digitization, to a program with more of a long term view. As a result, we are very much in flux when it comes to current policies and procedures.”

Others communicated their digital projects currently in process, “[The collection] will soon be on the Flickr Commons, probably before June 2009.”

A couple of respondents expressed anticipation for system or technological developments in the near future that would enable them to improve procedures. For example, one respondent said, “CARLI has most of the content we've digitized long term, and there hasn't been a way to track stats through them, but I have been told they are planning on making it possible.”

Another respondent added “Many tasks and policies (such as statistics gathering and preservation) will change in the near future as [CONTENTdm] software advances.”

Constraints and Limitations
These findings may or may not be transferable to other states depending on their digitization programs, an objective that is not within the framework of this study. A general consensus on the meaning of a digital collection haunted us from conceptualization through operationalization. A sample frame in which to compare our findings is a limitation of this study. While we counted at least 292 digital collections in February 2009, the sample reflected at least 224 collections in total among the respondents, of which two were unsure how many digital collections their organizations had. Technical questions and the length of the survey likely contributed to the missing data. Only the first three questions were required, however, if a respondent did not know an answer or could not easily attain it, they skipped the question, rather than check “I don’t know” some of the time.

The 21-day data collection period from March 16-April 6 may have been a constraint because of academic spring breaks. In addition, we were constrained by the time limit of our own semester. Triangulation with non-probability telephone and face-to-face interviews with key digital managers would enhance findings and provide richer qualitative data.

Lastly, when we began coding our qualitative data gathered through the “Other (please specify)” write-in option, as well as the last open-ended question, we discovered that one respondent wrote in answers on every allowable question, even where the appropriate answer had been provided on the questionnaire. While we appreciate such specificity and detail, this did skew our data numbers by a small percentage, as we opted to leave it intact and not alter the system-generated statistics.
Conclusion, Application and Trends

Libraries and organizations, foremost, are working to digitize collections that are valuable to their local communities. Overall, the results of this study show the bulk of activities centered on digital collections have generalizability in the State of Illinois. Future research can map this evolution, tracking how and what organizations are digitizing, marketing and sustaining in different regions over time. Ongoing education on developing technologies, such as OAI and xhtml/xml metadata tags, as well as marketing and comprehensive sustainability planning will benefit libraries and organizations that strive to make their collections accessible for generations to come.

At this time, image and text formats dominate the digital collections in the State of Illinois, with CARLI appearing to galvanize academic institutions. However, several organizations commented that they are presently working on oral histories. At the dawn of the 21st century, as Illinois’ libraries and organizations improve digital collections through best practices and standards, broaden in subject area and format, collaborate and market widely, they will transform into the interoperable digital libraries of the future. We are honored to have benchmarked this momentous day and time in history. Ultimately, our hope is that this study will function not only as a means to disseminate descriptive information about digital collections in Illinois, but also raise awareness about the multi-layered standards and practices of digital collection programs.
References


Hagar, C. (2009, January-April). Class notes LIS 769. Graduate School of Library and Information Science, Dominican University, River Forest, IL.


Appendices

A. Acknowledgements
   About the authors
B. Glossary of Terms
C. Questionnaire
D. List of Respondents
Appendix A

Acknowledgments

Special thanks to Lisa Barnhart at WebJunction Illinois, Karen Egan and Alyce Scott at the Illinois State Library, Nancy John, Associate Professor Emerita at the University Library of the University of Illinois at Chicago, Carole Medal, Karen Maki, Deb Huffman and Alissa Medows at the Gail Borden Public Library District, Kay Schlumpf at the North Suburban Library System, Dr. Michael Stephens at Dominican University. We are also indebted to our Research Methods professor, Dr. Chris Hagar, for her continued support of this project.

About the Authors

Miriam Lytle graduated with a B. A. in Anthropology from the University of Miami, Coral Gables, Florida. With a background in museums and international cultural travel, she began working at the Gail Borden Public Library District in Elgin, Illinois in 2004, where she is currently the Director of Communications and Program Coordinator. Miriam plans to graduate from Dominican University in December of 2009. m.lytle@sbcglobal.net

Marie M. Martino is a graduate student at Dominican University in River Forest, IL where she will complete her M.L.I.S. degree this May 2009. She also earned an M.F.A. in Moving Image from the University of Illinois at Chicago in 2007. She is interested in technology, user services, and instruction, as well as the space in which these areas overlap. marie@theotherlens.com

Justine Wagner-Mackow graduated from Northern Illinois University with B.F.A in Painting. For several years, she worked as a graphic designer. She is now a graduate student at the Graduate School for Library and Information Science at Dominican University in River Forest, IL. Justine will graduate in spring of 2010 with a focus in special collections. justine_moira@hotmail.com
Appendix B
A Descriptive Survey of Illinois’ Digital Collections Questionnaire
Designed by Miriam Lytle, Marie Martino and Justine Wagner-Mackow, Dominican University
Graduate School of Information and Library Science LIS 769 Spring 2009

Digital Collection Description
1. What is the name of your organization?

2. What is your organization type?
   - Academic
   - Corporate
   - Public
   - Special
   - Other (please specify)

3. How many digital collections does your organization have?

4. What is the general subject area(s) of your organization’s collection (per Illinois State Library 2009)? Please select all that apply.
   - African Americans
   - Agriculture
   - Arts and Architecture
   - Business/Industry/Manufacturing
   - Civil War
   - Coal
   - Disasters
   - Domestic/Community/Social Life
   - Education
   - Engineering and Construction
   - Genealogy
   - Government
   - Institutional History
   - Labor
   - Law
   - Local History
   - Maps
   - Medicine and Health Care
   - Military
   - Monuments and Memorials
   - Native Americans
   - Natural Resources
   - Newspapers
   - Oral History
   - Politics
   - Religion
   - Science and Technology
   - Transportation and Communication
   - Veterans
   - Women
   - World War I
   - World War II
   - Other (please specify)

5. What is your organization’s collection(s) significance? Please check all that apply.
   - International
   - Local
   - National
   - State
   - Other
6. What types of materials comprise your organization’s digital collection(s)? Please choose all that apply:
   - Books
   - Diaries
   - Drawings
   - Faculty publications
   - Glass negatives
   - Handwritten musical scores
   - Manuscripts
   - Maps
   - Material from your organization's archive
   - Material in the public domain published before 1923
   - Movie film
   - Newspapers with demonstrated historical significance
   - Oral histories
   - Personal journals
   - Photographs/Photographic negatives
   - Postcards
   - Posters
   - Prints
   - Sheet materials
   - Significant collections related to Illinois
   - Sound recordings
   - Tapes
   - Other (please specify)

7. Your organization’s digital collection(s) is comprised of the following formats (please choose all that apply):
   - Audio
   - Image
   - Text
   - Video
   - Other (please specify)

8. Who hosts your organization's digital collection(s)?
   - We host our own
   - An outside institution or agency
   - Other (please specify)

9. If an outside organization hosts your collection(s), please indicate which organization does so. Please check all that apply:
   - Barrington Area PL’s collections
   - CARLI (Consortium of Academic and Research Libraries in Illinois)
   - DePaul Universities
   - Digital Past (NSLS)
   - Illinois State Library (Illinois Digital Archives)
   - Illinois State University History, Milner Library
   - Lincoln Trail LS
   - University Library, University of Illinois at Urbana-Champaign
   - Upper Mississippi River Valley (includes Augustana College and other IL, not Iowa)
   - Other (please specify)
Selection Process and Planning
10. Does your organization have an official digital collection development policy?
   - Yes
   - No
   - I don't know

11. What best practices method does your organization employ when digitizing?
   - Illinois State Library (ILS)
   - Library of Congress
   - Cornell
   - University of Illinois at Urbana/Champaign (UIUC)
   - Institute of Museum and Library Services (IMLS)
   - None
   - Unknown
   - Other (please specify)

12. How is your organization’s digital collection(s) funded? Please check all that apply.
   - Internal Budget
   - LSTA grant
   - Private gift
   - Other (please specify)

Searchability and Sharing / User Access
13. What metadata standard does your organization utilize?
   - Dublin Core
   - Encoded Archival Description (EAD)
   - Federal Geographic Data Committee (FGDC)/ Content Standard for Digital Geospatial Metadata (CSDGM)
   - Global Information Locator Service (GILS)
   - Instructional Management Systems (IMS)
   - Machine-Readable Cataloging Record (MARC)
   - Text Encoding Initiative (TEI)
   - Visual Resources Association (VRA Core 3)
   - Other (please specify)

14. Does your collection comply with Open Archives Initiative (OAI) protocol to allow for metadata harvesting?
   - Yes
   - No
   - I don't know
15. Do you use metadata tags in your website’s xhtml document to make your collection more findable?
   - Yes
   - No
   - I don't know

16. Do you employ optical character recognition (OCR) to make text documents searchable?
   - Yes
   - No
   - I don't know

17. Does your organization market their digital collection(s)?
   - Yes
   - No
   - I don't know

**Marketing the Digital Collection**

18. If your organization does market their digital collection(s), how do they do it? Please check all that apply.
   - Blogs
   - Bookmarking
   - Facebook
   - Flickr
   - Genealogical Meetings
   - MySpace
   - On other organization's websites
   - On our organization's website
   - Outreach programs for students
   - Print
   - Second Life
   - Twitter
   - Wikipedia
   - Other (please specify)

19. Does your collection interface allow for user participation, such as tagging and/or commenting features?
   - Yes
   - No
   - I don't know

20. Does your organization work collaboratively with other organization(s)? Please check all that apply.

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
<th>I don’t know</th>
</tr>
</thead>
<tbody>
<tr>
<td>Museums</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other Libraries</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Universities/Colleges</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Businesses</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (please specify)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

21. If so, how do you work collaboratively?
- Collection procurement
- Funding
- Marketing
- Programs
- Scholarly research
- Not applicable to our organization
- I don't know
- Other (please specify)

**Maintenance and Sustainability**
22. How do you back up your collection? Please check all that apply.
- External storage, such as CD on site or off site
- External hard drive
- Internal computer storage
- LOCKSS (Lots of Copies Keeps Stuff Safe)
- OCLC Digital Archive Service
- Portico
- Tape backup on a server
- Other (please specify)

23. What does your organization have in place for the sustainability of the digital collection(s)? Please check all that apply.
- Plan
- Budget
- I don't know
- Other (please specify)

24. Does your organization keep usage statistics on your digital collection?
- Yes
- No
- I don't know

25. Is there anything else you would like us to know about your digital collection(s) that we have not already asked?

*Thank you for completing our questionnaire!*
*Miriam Lytle*
*Marie Martino*
*Justine Wagner-Mackow*
*Dominican University Graduate School of Library and Information Science*
Appendix C

Glossary of Terms

**Digital Collection** - defined as a collection of images, text, video or audio made available on the Internet or Intranet in any type of library or cultural institution

**Dublin Core** – may be used to describe a broad range of resource types (Northwestern University Library 2009)

**Encoded Archival Description (EAD)** – is used for encoding archival finding aids (Northwestern University Library 2009)

**Federal Geographic Data Committee (FGDC)/ Content Standard for Digital Geospatial Metadata (CSDGM)** – is designed to provide a common set of terminology and definitions for the documentation of digital geospatial data (Northwestern University Library 2009)

**Global Information Locator Service (GILS)** – can be used for all materials including books, datasets, objects, events, people while using GILS locator records (Northwestern University Library 2009)

**Instructional Management Systems (IMS)** – is used to describe electronic learning objects (Northwestern University Library 2009)

**LOCKSS (Lots of Copies Keeps Stuff Safe)** – based at Stanford University Libraries, is an international community initiative that provides libraries with digital preservation tools and support so that they can easily and inexpensively collect and preserve their own copies of authorized e-content (LOCKSS 2008)

**Machine-Readable Cataloging Record (MARC)** – was originally designed for the description of books and printed materials, however, it has now been expanded to allow for the description of other media (Northwestern University Library 2009)

**Metadata** – literally, data about data, or information that allows for resource identification, retrieval, and discovery (NISO 2004, p. 3)

**Open Archives Initiative (OAI)** – develops and promotes interoperability standards that aim to facilitate the efficient dissemination of content (Open Archives Initiative 2008)

**Optical Character Recognition (OCR)** – the identification of printed characters by means of equipment that scans them and reproduces them in electronic form (Oxford English Dictionary 2009)

**Text Encoding Initiative (TEI)** – is used to mark up and describe full-text materials (Northwestern University Library 2009)
**Visual Resources Association (VRA Core 3)** – is used to describe works of visual culture (e.g. painting, sculpture, building, performance, composition, literary work) as well as their surrogate images, such as digital, photomechanical and photographic (Northwestern University Library 2009)

**XHTML/XML Metadata Tags**—an element of hyper-text mark-up language that allows for metadata to be inserted into the XHTML/XML code that make up a webpage. This allows browsers to find and retrieve relevant pages more efficiently.
Appendix D. List of Respondents (n=44)

- Benedictine University
- Blackburn College
- Bloomingdale Public Library
- Blue Island Public Library
- Booth Library
- Brookens Library Archives/Special Collections, University of Illinois at Springfield
- The Career Center at University of Illinois Urbana-Champaign
- Center for the Study of Ethics in the Professions Library
- Chicago Ridge Public Library
- Columbia College Chicago Library Visual Resources
- DePaul University Library
- East-West University
- Eisenhower Public Library District
- Field Museum (Library Photo Archives)
- John M. Flaxman Library, School of the Art Institute of Chicago
- Paul V Galvin Library
- Graham Hospital School of Nursing Library
- Highland Community College
- Illinois College
- Illinois Wesleyan University
- Lincoln Trail Libraries System
- Melrose Park Public Library
- National-Louis University
- The Newberry Library
- Northeastern Illinois University
- Northern Illinois University
- Park Ridge Public Library
- Quincy University
- Rend Lake College
- Roosevelt University
- Roselle Public Library District
- Skokie Public Library
- St. Charles Public Library
- St. John's Hospital
- Southern Illinois University Edwardsville
- Southern Illinois University Edwardsville Library
- Western Illinois University
- University of Illinois at Chicago
- University of Illinois at Chicago Library
- University of Illinois at Urbana-Champaign
- University of St. Francis
- Anonymous (x 3)