Date : 01/07/2008

A Model of and Support for a Digital Preservation Infrastructure that Connects Individuals to Libraries

Andrea Japzon[1]
College of Information Science & Technology, Drexel University

Meeting: **84. Preservation and Conservation, (PAC), Information Technology, IFLA-CDNL Alliance for Bibliographic Standards (ICABS) and Law Libraries**

Simultaneous Interpretation: Not available

WORLD LIBRARY AND INFORMATION CONGRESS: 74TH IFLA GENERAL CONFERENCE AND COUNCIL
10-14 August 2008, Québec, Canada
http://www.ifla.org/IV/ifla74/index.htm

Introduction

Libraries and other memory organizations, such as museums and local historical societies, are very aware of the crisis in digital preservation and are taking steps to preserve our collective digital cultural heritage (LeFurgy, 2005; Ross & Hedstrom, 2005). In contrast, Marshall et al's (2006) research suggests that individual consumers are far less aware of the impermanent nature of their digital possessions, or to the extent that they are aware, feel disempowered to do anything about it. As a result, valuable representations of individuals' personal memories intended for future generations will be lost through ignorance and/or benign neglect (Yakel, 2004), and representations of family and social histories will be lost to what has been called the "digital dark ages" (Kuny, 1998). Many individuals are amassing large amounts of digital content because, like libraries, they have access to inexpensive and seemingly endless storage capability as well as to the high-powered computing needed to facilitate the creation and the downloading of digital content (Beagrie, 2005). However, the personal digital information environment offers limited infrastructure for content organization and

[1] Andrea Japzon, College of Information Science & Technology, Drexel University

This research was supported by the Laura Bush 21st Century Librarian Fellowship provide by the Institute of Museum and Library Services. Special thanks to Denise Agosto, Jennifer Burke, and Nadia Nasr for their help with the survey and paper.

Correspondence concerning this article should be addressed to the author at College of Information Science & Technology, Drexel University, 3141 Chestnut Street, Philadelphia, PA 19104. E-mail: Andrea.Japzon@drexel.edu

preservation, so the likelihood that individuals will lose valuable representations of personal memories is very real (Jones, 2007).

For libraries and other memory organizations to create a sustainable infrastructure for digital preservation they must collaborate with each other, create interoperable systems, and develop standards for creation and storage (Bradley, 2007; Mason, 2007). This paper explores the possibility of a preservation infrastructure that includes the preservation of personal digital information, the type of information that one keeps for one's own purposes, (e.g. photographs, letters, emails, music, web pages, diaries, and videos). Who will individuals collaborate with to preserve their own personal information, how will individuals know what the best practices for storage and file formats are, and in what archival infrastructure will such valuable personal digital information reside? In corporate and academic library environments institutional repositories are being developed to meet the long term digital preservation needs of scholars and corporate researchers (Lynch, 2003). Could public libraries and other local memory organizations work with their constituents to create community repositories for the preservation of personal information, which also contributes to social and cultural histories?

There is an evolving infrastructure of personal digital information storage and implications for long-term preservation. Personal information storage has been gradually moving away from a client-side approach utilizing hard drives and local installations of software applications toward the web-based storage and services models offered by private companies such as Google and Flickr (Carr, 2008). Libraries and archives have a long history of educating individuals on how to preserve their personal analog information, but these efforts have not yet evolved to include digital media. Should this tradition of preservation education in the digital era be left to Google, or should information professionals take this opportunity to reassert their skill, credibility, and leadership in this domain? A model of a preservation infrastructure is proposed that encompasses libraries and other memory institutions interconnecting with individuals through a multi-faceted approach involving education, technology, community, and shared expertise is proposed. Additionally, results are presented from an exploratory investigation of current practices by memory organizations to include community members in the creation of digital collections.

A Conceptual Model

For many reasons, libraries and other memory organizations are excellent sources for the provision of education to individuals in the practices of digital preservation. First, libraries are either gaining or creating state-of-the-art knowledge regarding digital preservation as they create and maintain their own digital collections. Second, libraries are a vital part of communities and of the information economy and in such have the power to create connections between the two. Third, libraries already serve the role of educating users through information literacy programs. Lastly, libraries are creating digital collections of local significance that could easily attract the involvement of community members.

Further, community members are excellent resources for the creation of community collections. They provide knowledge of local history, customs and events and an interest in supporting community organizations. From the museum science community, Russo and Watkins (2007) describe "community cocreation" as cultural institutions and communities working together to create digital content, each benefiting and learning from the other's expertise and experience. Library and information professionals working through their organizations provide a technical infrastructure and contribute technical expertise in collection development and maintenance. Community members provide the knowledge that supports content development and contribute personal digital information and physical artifacts from their own collections. Figure 1 illustrates the basic structure of a community repository model.

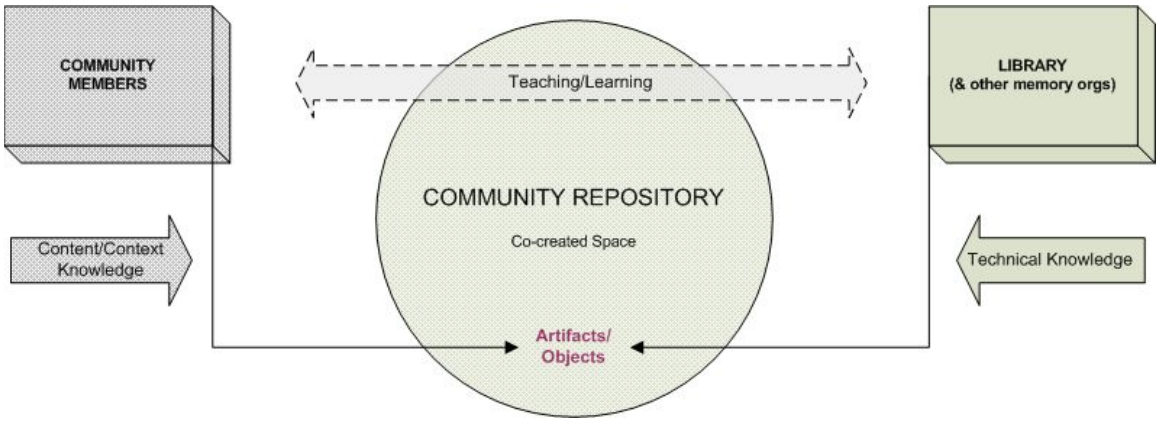

Figure 1. Model of Community Repository

A community repository would provide the space for collaboration, creating a deep collection of local history reflecting the knowledge of community members. This collaboration would bring together holdings from library and personal collections along with the expertise of library and information science professionals. Community members would not only contribute personal artifacts and information to co-created collections but also the cultural and historical context for the items being collected. Their personal and cultural experiences would provide the narrative that gives the collection meaning. The online exhibits of the Olympic Peninsula Community Museum provide an excellent example of this.[i] The exhibits related to Peninsula-based tribes were created in participation with those tribal communities. Other examples of digital collections exist that support the notion that libraries are aware of the value of community member knowledge in creating digital collections.[ii] However, are community members benefiting from the technical knowledge from library and information science professionals? The collection development process can then ideally become one of exploration rather than exploitation. The results from the exploratory study discussed below indicate this ideal does not yet exist as the exchange of knowledge is not equal.

Collection creation, maintenance, and preservation (physical or digital), are knowledge intensive processes. Through the co-creation of digital collections, librarians and other information workers could provide community members with hands-on context-driven learning experiences. Community members could learn first hand the value of appraising content, selecting formats, applying metadata, and the implications of technology infrastructure for information storage and access. They could also provide community members with a trusted source to seek information related to digital information management and preservation. Libraries and archives have a tradition of helping individuals care for personal tangible information such as letter and photographs. Generally, libraries have connected with users by providing access to computers and networked proprietary resources for personal use, but libraries have not succeeded nearly as well at connecting with individuals in their personal computing space. Creating community repositories is a means for libraries to connect to individuals' personal computing space in a way that plays to libraries strengths rather than competing with such enterprises as Google, Facebook, Amazon, or Flickr.

Exploratory Study

A survey was designed to gather input from digital collections professional who are interested in creating digital collections for community building and information literacy instruction. The survey was distributed through relevant listservs subscribed to by digital collections professionals including: DIGLIB, DigiStates, Archives & Archivists, IMAGELIB, and DIGPRES. The survey was created using Vovici web survey software with distribution and collection of responses online. 104 responses were collected from July 2007 to October 2007. Response rate was low considering there is a combined total of over 8,000 subscribers to these listservs. While survey results are limited, the responses offer some insight into the ways community members are currently included in the creation of digital collections.

The survey questions focused on the circumstances in which community members participated in digital collections including who initiated the projects, what role community members played, whether there was inter-agency collaboration, whether the projects included information literacy components, details about collection policies, and details about the collection and organization of metadata. Multiple choice and open-ended questions were employed. The data were analyzed using descriptive statistics.

Analysis

Survey participation came mostly from individuals working in academic libraries and from those located in the United States (78.8%). Table 1 illustrates the diversity of institutions represented in the study. While the majority of respondents working with digital collections are located in academic libraries, the diversity of institution types represented indicates how widespread digital collection creation is throughout cultural institutions. These findings offer limited support for the notion that academic library patrons have greater access to digital collections and the attendant professional expertise than do public library users. It also suggests that more resources and services need to be made available to public library users in this regard.

Table 1
Percent of participants by type of institution

Type of Institution	*Percent*
Academic Library	50.5
Federal Library	1.0
Public Library	8.1
Special Library	3.0
Local History Society	2.0
Art Museum	6.1
Science Museum	1.0
Social History Museum	4.0
Other (archives, academic departments, government agencies, associations, and organizations)	24.2

The range of community involvement reported by those responsible for digital collections can be seen in Table 2. The survey provides evidence that community members are being included in the creation of digital collections. Sixty-five percent of the respondents included community members in digital collections projects and an even greater seventy percent expressly included items at the request of patrons. These statistics alone indicate that community member interaction with library digital collections is not one of solely passive use, but also of active creation. While it is good to note that nearly half of organizational respondents have collection development policies regarding digital materials, only eleven percent of those policies specifically include "involving community members". There is additional room for community engagement and learning as an information literacy component was included only 29% of the time.

Table 2
Percent of participants who have included community members

Overview of Community Involvement	*Percent*
Included Community Members in Digital Collection Creation	65.0
-Included an Information Literacy Component	28.8
Have Digital Collections Development Policy	45.5
-Included Community Member Involvement	10.6
Included items in Digital Collection as a Result of Patron Request	70.1

In response to why survey participants included community members, one participant offered this insightful view of communities and collections: "In order to provide a collection of works the community members must be involved to make the collection a success. Digital collections are not just technical information sources they are social information sources as well." The rationales seen in Table 3 support this emphasis on

community building and cooperation. The open-ended "other" category provided interesting additional reasons for inclusion that were more practical: free labor and access to cultural and local history knowledge. Offering information literacy instruction was a low motivator for inclusion at less than nine percent.

Table 3
Percent of participants by reasons for inclusion of community members

Reason for Including Community Members	*Percent*
Community building	31.7
Information literacy instruction	8.7
Community members requested to be involved	24.0
Cooperation with external agency	15.4
Cooperation with an internal agency	5.8
Other (provided for free the following: work, expertise, metadata, knowledge, materials, information, archives, financial support, object identification, scanning, appraisal, oral history, writing skills; promoted digital projects; and added legitimacy to cultural representations)	27.9

As seen in Table 4, the main role of community member involvement was digital collection creation: from item selection and description to creation (scanning). To a lesser extent community members were also involved in supporting collections through promotion and fundraising. It is encouraging to learn that digital information and artifacts from personal collections are being included in library digital collections. This finding supports the idea that a cultural heritage infrastructure can be extended to include information and artifacts that represent the lives of individual community members. Connections between libraries and individuals can be made through an interest in community and cultural history and the objects that represent that history. This shared interest creates a communication vehicle through which other issues related to digital collections creation and preservation can evolve.

Table 4

Percent of participants by type of community member involvement

Type of Community Member Involvement	*Percent*
Contributed personal artifacts to be digitized	30.8
Contributed personal digital objects	21.2
Recommended artifacts to be digitized	26.9
Recommended digital objects	21.2
Helped with scanning	20.2
Helped with the creation of metadata	32.7
Helped to promote collections	27.9
Other (identified artifacts, raised funds, and wrote letters recommending project)	10.6

While collection scope, selection, appraisal, and the use of metadata and scanners are not typically thought of as subjects for information literacy instruction, they are important to creating and maintaining collections of information. If individuals are going to handle long term preservation of personal information, they will need to learn to deal with their personal digital information items in terms of collections rather than simply stores of information located either on personal digital devices or in web-based storage sites. Librarians, archivists, and other information professionals possess a wealth of knowledge that would aid in this learning process. Through the creation of community repositories, individuals can learn from them, these new types of information literacy skills in a context-driven learning environment.

The survey responses indicate an emphasis on instruction related to collection, scope, selection, and descriptive metadata and almost no emphasis on preservation metadata (see Table 5). Without knowing more, the assumption would be that preservation metadata is not an integral part of the digital collections creation process. Day and Alemneh et al (1997, 2002) assert that preservation metadata has been an after-thought in creation of digital collections when, in truth, the capture of preservation metadata needs to begin with the creation of each digital object. The technical information needed to support each digital object should be part of its creation, because without knowledge of the technical infrastructure the object will not remain accessible over the long term. Initially, digital collection creators gave most thought to describing content to promote content discovery. As digital collections and digital content acquisitions continue to grow in size institutions are beginning to plan and budget for the preservation of the digital collections and content they are acquiring (Searle & Thompson, 2003). Institutions are beginning to address the fact that digital content can not be acquired without the simultaneous consideration of preserving that content. This is an important realization that needs to be shared with community members as they acquire personal digital information.

Table 5

Percent of participants by type of information literacy component

Information Literacy Component	*Percent*
Collection Scope	18.3
Selection	22.1
Appraisal	8.7
Broadly representing content for re-use	12.5
Descriptive metadata	19.2
Structural metadata	7.7
Preservation metadata	4.8
Scanning	12.5
Other (general standards for digitization, general knowledge and familiarity with using computers, and the creation of transcriptions)	3.8

The most common element among existing digital collection policy statements addresses format recommendations, which is also an important consideration for long term preservation (see Table 6). Elements specific to technical aspects of digital preservation practice such as migration (17%), and refreshing data (12%) received little consideration. Yet it is these technical elements which may most determine the overall success of digital preservation, and prove the toughest barriers to overcome. In this small sample, digital preservation practices and community involvement were not among the most prevalent policy elements. This is unfortunate, particularly if this sample is reflective of other policies, because these two elements are necessities for digital collections sustainably.

Table 6

Percent of participants by type of digital collections policy element

Digital Collections Policy Elements	*Percent*
Migration	17.3
Refreshing data	11.5
Rules for metadata	31.7
Scope	26.0
Format recommendations	31.7
Community member involvement	10.6
Other (intellectual property rights, potential for collaboration, funding support, sustainability for materials, topics of focus, digitization for preservation, and content and selection criteria)	5.8

Conclusion

The community repository model presented here calls for libraries and other organizations to include their community members in the infrastructure that supports their digital cultural heritage initiatives. Include community members through the sharing of technology for both consumption and creation of digital content. Include them by having policies that acknowledge and prioritize their importance to digital collections that reflect local culture and history. And include them in the learning process by sharing collection development expertise and best practices for digital preservation.

While limited, the survey does offer some encouraging support for the future development of community repositories. Foremost, personal information and artifacts of community members are already being incorporated into digital collections. This indicates that the long standing connections that libraries have with their communities is reflected in their digital initiatives. Also of significance is the fact that community members are being taught how to create digital collections. This new knowledge can then in turn be applied to their own personal information collections.

The preservation of cultural and historical digital information is everyone's problem, from governments to cultural institutions to individuals. The solution to this collective problem is to be found in research and cooperation. The community repository model includes individual consumers in the solution in a way that benefits both cultural heritage and the individuals' own personal collections of information. Further research is needed on the role of community members in library digital collections, on the importance of digital preservation to digital collection creation, and on the connections between the expertise of information professionals and the knowledge needed to management personal information collections over the long term. This exploratory study indicates that there is potential for further research and development in these areas.

References

Alemneh, D. G., Hastings, S. K., & Hartman, C. N. (2002). A metadata approach to Preservation of digital resources: The University of North Texas Libraries' experience. *firstmonday*, (7)8. Retrieved from http://www.firstmonday.org/issues/issue7_8/alemneh/index.html on November 13, 2006.

Beagrie, N. (2005, June). Plenty of room at the bottom? Personal digital libraries and Collections. *D-Lib Magazine, 11*(6). Retrieved from http://www.dlib.org/dlib/june05/beagrie/06beagrie.html on February 3, 2007.

Bradley, K. (2007). Defining digital sustainability. *Library Trends, 56*(1), 148-163.

Carr, N. (2008). *The Big Switch: Rewiring the World from Edison to Google.* New York: W. W. Norton & Company.

Day, M. (1997). Extending metadata for digital preservation. *Ariadne, The Web Version.* Retrieved from http://www.rlg.org/preserv/joint/day.html on April 9, 2007.

Kuny, T. (1998). The Digital Dark Ages? Challenges in the preservation of electronic information. *International Preservation News,* No. 17. Retrieved on October 22, 2006 from http://www.ifla.org/VI/4/news/17-98.htm#2.

LeFurgy, W. (2005, Summer). Building preservation partnerships: The Library of Congress National Digital Information Infrastructure and Preservation Program. *Library Trends, 54* (1), 163-72.

Lynch, C. (2003). Institutional repositories: Essential infrastructure for scholarship in the digital age. *portal: Libraries and the Academy*, 3(2), 327-336.

Marshall, C., Bly, S., & Brun-Cottan, F. (2006). The long term fate of our digital belongings: Towards a service model for personal archives. In Proceedings of IS&T Archiving 2006, Ottawa, Canada, May 23-26, 2006. Springfield, VA: Society for Imaging Science and Technology, pp. 25-30.

Mason, I. (2007). Cultural information standards – Political territory and rich rewards. In F. Cameron & S. Kenderdine (Eds.), *Theorizing Digital Cultural Heritage* (pp. 223-244). Cambridge, MA: MIT Press.

Ross, S. & Hedstrom, M. (2005). Preservation research and sustainable digital libraries. *International Journal on Digital Libraries, (5),* 317-324.

Russo, A. & Watkins, J. (2007). Digital cultural communication: Audience and Remediation. F. Cameron & S. Kenderdine (Eds.), *Theorizing Digital Cultural Heritage* (pp. 149-164). Cambridge, MA: MIT Press.

Searle, S. & Thompson, D. (2003, April). Preservation metadata. *D-Lib Magazine, (9)*4. Retrieved from http://www.dlib.org/dlib/april03/thompson/04thompson.html on April 2, 2007.

Yakel, E. (2004, October). Seeking Information, seeking connections, seeking meaning: genealogists and family historians. *Information Research, 10*(1). Retrieved from http://informationr.net/ir/10-1/paper205.html on February 11, 2007.

[i] http://content.lib.washington.edu/cmpweb/project/groups.html
[ii] http://c9.mdch.org/
http://www.mdch.org/collection.aspx?id=17308
http://911digitalarchive.org/
http://www.storycorps.net/