

# Report on the Minnesota Digital Library Conference

August 2001

## Introduction

At the invitation of MAGNOLIA, the Minnesota Academic Group for New Opportunities in Library and Information Access<sup>1</sup>, and with the support of an LSTA grant, twenty-five librarians, archivists, and technologists (list of participants in Appendix A) representing a broad spectrum of cultural heritage and memory institutions, came together 14-16 August in Monticello, Minnesota, to discuss the possibility and initiate planning for a coordinated collaborative digital library project. The summit set ambitious goals that required the participants' talents, enthusiasm, and energy to fulfill.

- ?? Establish a common definition, purpose, scope and goals for a Minnesota Digital Library (MDL) that accommodates the collections and requirements of a wide range of institutions and the needs and expectations of a wide variety of user communities.
- ?? Explore potential sources of external funding to support a widely distributed collaborative digital library project.
- ?? Provide a foundation for a shared policy standard for baseline metadata information
- ?? Identify technical issues that would confront content and access issues for a digital collaborative
- ?? Identify parameters for participation and examine possible governance structures for a collaborative project.

The idea for a collaborative digital library is not unique to Minnesota. Several states, notably Colorado,<sup>2</sup> Missouri,<sup>3</sup> Washington,<sup>4</sup> North Carolina,<sup>5</sup> Kentucky,<sup>6</sup> Georgia,<sup>7</sup> and California,<sup>8</sup> have already undertaken digital library projects with varying levels of collaboration among participants and with widely divergent project definitions. The one

---

<sup>1</sup> MAGNOLIA is the appellation adopted by a group of academic library directors, representing the libraries at University of Minnesota, St. Cloud State University, Minnesota State University, Mankato, and Winona State University, who along with the director of MINITEX occasionally meet to discuss topics of mutual interest. The group initially gathered to explore ways in which their institutions could contribute to the long-term value of Minnesota's educational, cultural, and historical resources. MAGNOLIA participants see the development of a digital library for Minnesota as a project through which they can apply the experience and expertise of their combined institutions.

<sup>2</sup> Colorado Digitization Project (<http://coloradodigital.coalliance.org/>; accessed 30 August 2001).

<sup>3</sup> Virtually Missouri (<http://www.virtuallymissouri.org/>; accessed 30 August 2001).

<sup>4</sup> Washington State Digital Library (<http://digitalwa.statelib.wa.gov>; accessed 30 August 2001).

<sup>5</sup> North Carolina ECHO (<http://www.ncecho.org/>; accessed 30 August 2001).

<sup>6</sup> Kentucky Virtual Library (<http://www.kyvl.org/>; accessed 30 August 2001).

<sup>7</sup> GALILEO (<http://neptune.libs.uga.edu>; accessed 30 August 2001).

<sup>8</sup> California Digital Library (<http://www.cdlib.org>; accessed 30 August 2001); includes the Online Archive of California.

characteristic common to all of these projects is the expansion of access through digitization for rare or remote resources to share the cultural heritage and experience of a state with people. The Minnesota initiative has the opportunity to learn from and build upon the experience of earlier projects.

Liz Bishoff and Nancy Allen of the Colorado Digitization Project assisted in preparing the conference agenda. Nancy Allen presented a keynote overview of issues<sup>9</sup> confronting a digitization project and the benefits of multi-institution collaboration. During the first two days of the conference, Ms. Allen provided insightful comments in both general and small-group discussions.

### **Vision and Mission**

Early in the discussion, conference participants became aware of the enormity of both the MDL project and the opportunity inherent in a digitization collaborative. The scope of MDL collections is unlimited, with potential content drawn from the diverse collections of Minnesota's archives, libraries (including public, academic, corporate, and state agency), historical societies, and museums. Senator Steve Kelly, a participant on the final day of the conference, went so far as to prompt participants to consider ways to include cultural activities, like traditional construction of a birch bark canoe, as components with the digital library. Access to the collections presents opportunities to expand the scope of the project beyond simple search and display to exhibition, interpretation, and curriculum integration. Conference participants discussed these issues in considerable detail over the course of the conference.

From the beginning, participants decided to distinguish the MDL initiative as a digitization project distinct from the Electronic Library for Minnesota (ELM)<sup>10</sup> and as an access project distinct from but linked to MnLINK.<sup>11</sup> Many potential MDL participants are not now participants in the MnLINK integrated library system; indeed, few maintain automated catalogs of resources in their collections. It is not the intent of the MDL project to fund or undertake library automation projects, although such projects would be encouraged. For the time being, conference participants agreed to adopt the Digital Library Federation definition of a digital library:

---

<sup>9</sup> Liz Bishoff is Director of the Colorado Digitization Project (CDP); Nancy Allen is Dean and Director of Libraries at the University of Denver and principle investigator for the CDP. Ms. Allen's PowerPoint presentation for the conference is available online at <http://lrs.stcloudstate.edu/mdl/allenpresentation.html>.

<sup>10</sup> ELM (<http://www.minitex.umn.edu/eit/infosheet.asp>; accessed 30 August 2001) is a Minitex coordinated, legislature-funded initiative to provide access to subscription-based commercial electronic information resources to Minnesotans through their local or regional public library, school library, or academic library. GALILEO and the Kentucky Virtual Library include commercial databases as a major component of their projects.

<sup>11</sup> Minnesota Library Information Network (<http://www.mnlink.org/>; accessed 30 August 2001) is a statewide initiative to share an integrated library system among academic libraries, public libraries, state agency libraries, and school libraries.

“[An organization] that provides the resources, including the specialized staff, to select, structure, offer intellectual access to, interpret, distribute, preserve the integrity of, and ensure the persistence over time of collections of digital works so that they are readily and economically available for use by a defined community or set of communities.”<sup>12</sup>

At the conclusion the conference, participants agreed upon the following mission statement:

**The archives, historical societies, libraries, and museums of Minnesota will collaboratively build a coordinated access to Minnesota’s unique resources and special collections in digital format for use by the people of Minnesota and, through the Internet, the people of the world.**

### **Themes and Issues**

Among the initial ideas underlying the MAGNOLIA grant application and discussed by conference participants were the need to assist smaller local and regional libraries, museums, and historical societies to undertake digitization of portions of their collections, to facilitate and enhance user access to distributed collections through a common portal or metasearch capability, and to develop a collaborative structure that would recognize the value from each and to each participant. Conference participants explored these ideas through four discussion groups developed by the conference planners:

- ?? participation, policy and planning,
- ?? standards and digital content creation,
- ?? audience presence, and
- ?? collections.

The following sections of this report are based on these discussion groups and their exploration of important themes and issues in developing a Minnesota Digital Library.

### **Participation, Policy, and Planning**<sup>13</sup>

Conference participants envision the MDL as a collaboration among cultural heritage or memory institutions, with potential participation from every archive, library, museum, and historical society in Minnesota. To encourage broad participation, they suggested there should be minimal impediments and maximum benefits to participation. Conference planners selected participants to represent diverse institutions and collections; planners and participants suggested other institutions, for example the Legislative

<sup>12</sup> See Introduction DLF Annual Report for 1998-1999 (<http://www.diglib.org/ar9899p1.html>; accessed 30 August 2001).

<sup>13</sup> Members of the participation discussion: Kristi Tornquist, Mike Kathman, Joyce Swonger, Eileen McCormick, Christine Clements, Sherry Sweetman, and Steve Harsin.

Reference Library, that should be invited to join further development discussions. In addition, there was some agreement that representatives from various user communities, including scholars, genealogists, and elementary-secondary instructors, could add substance to discussions of some working groups. Many conference participants were unable to commit their institutions to formal participation in the MDL but agreed nonetheless to continue participating in development discussions.

Governance is a more difficult issue, eventually depending on both funding and the form of the collaboration. The long-term success of the MDL will be assured only through collaboration, bringing the experience and expertise of diverse institutions together to create and sustain the project. In order to do this, however, MDL participants will need to adapt to joint planning and comprehensive strategies that have an impact on home institutions; likewise, MDL policies and structures will need to be flexible to accommodate many of the concerns of participating institutions. Among the more difficult issues to resolve will be ownership and authority.<sup>14</sup>

While there is a strong desire to retain an informal structure for the present, conference participants recognize the need to create a more formal collaborative coordination structure in the future. Based somewhat on discussion groups during the conference, participants agreed to establish four working groups to continue planning and prepare the next grant application:

- ?? Policy and planning<sup>15</sup>
- ?? Standards and training<sup>16</sup>
- ?? Audiences and presence<sup>17</sup>
- ?? Collections<sup>18</sup>

The interim chairs of the working groups will constitute a temporary executive committee until organizations commit to participation and planners propose and implement a more formal. This minimal structure, operating with support from MAGNOLIA, will coordinate development in the short term and provide a structure from which to submit grant proposals for further project development.

The group proposed a formal structure that would establish long-range project production planning, maintain financial oversight, set policies, and promote collaboration, but they deferred to a later date pending further discussion. The proposal envisioned a steering committee with representatives drawn from different types of participating institutions,

---

<sup>14</sup> See Betsy Wilson *The Lone Ranger is dead: Success today demands collaboration* (<http://www.ala.org/acrl.wilson.html>; accessed 30 August 2001) and Diane H. Sonnenwald, et al. (1999) *Collaboration Services in a Participatory Digital Library: An Emerging Design*, from the *Proceedings of Third International Conference on Conception of Library and Information Sciences*, 23-26 May 1999.

<sup>15</sup> Policy and planning working group: Kristi Tornquist (chair), Keith Ewing, Mike Kathman, Eileen McCormick, Joan Roca, and Christine Clement

<sup>16</sup> Standards and training working group: Charles Thomas (chair), Jason Roy, Eileen Quam, Steve Harsin, Robin Dowden, Joyce Swonger, and Scott Sayre

<sup>17</sup> Audiences and presence working group: John Butler (chair), Chris Olson, Steve Harsin, Marian Rengel, Scott Sayre, Sherry Sweetman, and Keith Ewing

<sup>18</sup> Collections working group: Daardi Sizemore (chair), Sue Ellingwood, Tom Malefatto, Mark Reidell, Betsy Williams, and Nancy Broughton

with one representative from each of the following: Minnesota Historical Society, University of Minnesota Libraries, Minnesota State Colleges and Universities libraries, county historical societies, museums, archives, k-12 education, state agency libraries, and public libraries; two additional at large representatives would be selected and appointed by MAGNOLIA. The tenor of the discussion was to support collaboration and collective actions yet retain considerable independence for participating agencies, avoiding as much as possible the difficulties that have beset the MnLINK project.<sup>19</sup>

Funding a project of this scope will be complex. Federal grants from Institute for Museum and Library Services (IMLS) may assist in developing initial collaborative projects, which in turn will establish a foundation for future development; IMLS grants may also be available to developing unique applications that could be adopted by other digitization projects. Private foundation grants, while very competitive, are also available, especially from regional foundations. There was also discussion of corporate sponsorship of the MDL project, especially for corporations whose archives contain resources that could be integrated into MDL collections. Setting user fees for access to digital resources was discussed but ultimately rejected. Participants agreed to leave the discussion of access fees to a later date, especially for high-resolution images for download or for alternative media. There is the possibility that participation or membership fees could support a portion of the project, but such fees should be minimized to encourage rather than inhibit participation. This would not preclude fees should a small collection choose to house their digital objects on an MDL server. Finally, there is the possibility of going to the Minnesota legislature to secure funding. Senator Kelly noted that this is the type of project that the legislature supports, but cautioned the group to be aware that funding the MDL project may diminish funds elsewhere in the state budget that may be of equal import to MDL participants. In the absence of a formal governance structure for MDL, MAGNOLIA members will have to assume responsibility for resolving the funding dilemma in the short term.

Providing mini-grants for digitization of specific resources or collections should be incorporated into an early grant application. This will provide a focus for developing digitization and indexing routines, training requirements, and collaboration processes as well as heighten public awareness of the project. How to select resources for digitization is discussed elsewhere in this report. A process for funding mini-grants (application, review, and reward) needs to be established. There is much planning to be done, however, before reaching this point.

## **Standards and Digitization<sup>20</sup>**

### **Metadata Standards**

---

<sup>19</sup> MnLINK is attempting to bring together different types of libraries with varying collections and missions. There are several difficulties in implementing an appropriate governance and funding structure.

<sup>20</sup> Participants in the standards and digitization discussion: Charles Thomas, Eileen Quam, Jason Roy, Tom Malefatto, and Daardi Sizemore.

Whether digital library planning encompasses a global, distributed effort or a local initiative, choosing a suitable metadata standard is the greatest determinant of success.<sup>21</sup> With several probable MDL participants already involved in well-established digitization initiatives and using widely accepted descriptive metadata standards<sup>22</sup>, and with diverse institutions and collections as potential participants, selecting an appropriate single descriptive metadata scheme could inhibit participation, diversity, and the ultimate success of the MDL project. Clearly, adopting a single existing metadata scheme, such as the widely used Dublin Core Metadata Initiative scheme, would be best for interoperability and data exchange but is probably insufficient for the heterogeneous MDL environment. Yet, for a large-scale project such as that envisioned for the MDL, the limitations and deficiencies of any single descriptive metadata scheme become apparent. Local development of a complex, integrated descriptive metadata scheme that anticipates and accommodates all potential applications would be prohibitively expensive and ultimately difficult to adapt to particular needs, thus inhibiting participation in the project. Obviously, the less metadata required, the simpler the search and retrieval; more complex metadata can result in robust search capabilities. Conference participants recognized that further discussion is necessary to establish an optimal solution for the MDL project.

Members of the discussion group encouraged selecting a core set of necessary descriptors common to all metadata schemes or with crosswalks between schemes while supporting optional elements that could be based on any scheme. Thus there would be a range of descriptive metadata from those required by the MDL to optional descriptors dependent upon local needs. Similar flexibility could apply to administrative metadata. The standards discussion group concluded that to the degree possible and reasonable, MDL should use existing metadata standards, which are necessary for adequate description and precision retrieval, in order to simplify implementation and to support national metadata harvesters and search engines.<sup>23</sup> This also requires attention to and participation in continuing international developments with metadata standards and implementations.

As another means of encouraging participation, the group suggested that the MDL could provide a template or multiple templates for creating metadata<sup>24</sup> or adding locations (holdings) to existing metadata in a shared catalog database. This central shared catalog will serve three purposes: to provide a gateway for multiple users to distributed digital objects, to provide a framework for database enhancement (see Public Presence below), and to create a public image for the project. This central catalog would complement rather than replace existing metadata repositories, with individual collection metadata

---

<sup>21</sup> Charles F. Thomas (2001) Choosing Digital Library Metadata Standards. Unpublished paper distributed at the MDL conference. Much of the following combines the group discussion notes with Thomas's paper.

<sup>22</sup> See Metadata Resources (<http://www.mnhs.org/preserve/records/meta.html>; accessed 30 August 2001) for links to a variety of metadata schemes.

<sup>23</sup> See, for example, the Open Archives Initiative (<http://www.openarchives.org>; accessed 30 August 2001) and the Research Library Group's Cultural Materials Initiative (<http://www.rlg.org/culturalres/>; accessed 30 August 2001).

<sup>24</sup> See, for example, the Dublin Core Metatag Generator (<http://bridges.state.mn.us/taggen.html>; accessed 30 August 2001) developed by the Foundations Project in partnership with Hiawatha Island Software.

residing in both locations. A simple means of uploading existing metadata to a central repository would need to be developed.

To simplify the adoption of a metadata standard across the range of possible MDL participants, the discussion group realized it might be wise to start with a metadata scheme for a specific domain of digital objects, such as still images, before adding complexity to accommodate heterogeneous collections. Deciding where to begin, that is, what type of material and what collection/s to digitize would facilitate selection of a foundation metadata scheme.

The goal for MDL metadata should be to describe a digital object once and use the description many times. This will require a means within the database to show multiple locations (holdings) of the same digital object. While there is no way to avoid some duplication of metadata records, as experience with OCLC MARC records shows, MDL participants should attempt to minimize their occurrence in order to lessen user confusion and frustration.

### **Digitization Standards**

Digitization involves several challenges<sup>25</sup> beyond metadata implementation, among those identified by conference participants are

- ?? Adopting and adapting widely accepted digital content creation principles and standards,<sup>26</sup>
- ?? Acquiring the appropriate hardware to capture or create digital objects and to store and disseminate those objects,
- ?? Identifying appropriate resources to digitize<sup>27</sup> (based on need to expand access, need to integrate into an online exhibition or curriculum enhancement project, or need to promote preservation of original object), and
- ?? Sustaining digital objects and access to them over time.

Every digitization project also struggles with the need to add value to digital objects and overcoming the tension between access and preservation. These challenges will not go away in a collaborative environment, and for individual institutions may become more

---

<sup>25</sup> The Library of Congress National Digital Library Program identifies several challenges. See Challenges to Building an Effective Digital Library (<http://memory.loc.gov/ammem/dli2/html/cbedl.html>; accessed 30 August 2001).

<sup>26</sup> See, for example, *NARA Guidelines for Digitizing Archival Materials for Electronic Access* (<http://www.nara.gov/nara/eap/eapspec.html>; accessed 30 August 2001), *Building Digital Collections: Technical Information about American Memory Collections* (<http://memory.loc.gov/ammem/techdocs/digcols.html>; accessed 30 August 2001) and Howard Besser Procedures and Practices for Scanning (<http://sunsite.berkeley.edu/Imaging/Databases/Scanning/>; accessed 30 August 2001). The Colorado Digitization Project has adapted Besser in General Guidelines for Scanning (<http://www.coloradodigital.coalition.org/scanning.html>; accessed 30 August 2001).

<sup>27</sup> See Paul Ayris (1999) Guidance for selecting materials for digitization (<http://www.rlg.org/preserv/joint/ayris.html>; accessed 30 August 2001) from the Joint RLG and NPO Preservation Conference. For a sample selection policy, see Columbia University Library Selection Criteria for Digital Imaging (<http://www.cc.columbia.edu/cu/libraries/digital/criteria.html>; accessed 30 August 2001)

difficult, but collective solutions may be more durable and rewarding for all MDL participants.

Conference participants could not have been stronger in their support for adhering to national standards for capture and storage of digital images that maintain fidelity to the original, especially photographs, prints and artwork, or images of text. Typically a single object (e.g., photograph) will need three files (thumbnail, access, and master or archive) requiring three different file formats; display of each file would be associated with a different level of metadata completeness. Image objects, however, present problems for visually impaired users and conference participants were eager not to exclude this audience.

Textual objects can be more complex. While the Text Encoding Initiative (TEI) is a widely used metadata scheme for digital textual resources, there are no national standards for digital reproduction of books or serials. In July 2001, the DLF issued a draft benchmark<sup>28</sup> that may provide an appropriate standard in the future. The JSTOR model,<sup>29</sup> using OCR to scan the fulltext to support keyword searching and page images for display, offers a potential solution, despite relatively higher production costs, that should be explored further. JSTOR does not use the OCR text to support access for visually impaired users with speech software; this should be included in future MDL discussion of this issue.

In the absence of firmly established standards for all formats of media, compiling and disseminating a set of best practices provides a viable solution. Best practices identify the optimal implementation and uses of particular technologies, processes, and storage formats. Conference participants recommended that these practices be documented and readily accessible via the Web or other media to all potential participants in the MDL.

Distributed digital objects require a consistent and coordinated naming scheme for all digital objects. Implementing a naming service would assure unique names for digital objects within MDL. Such an undertaking may create problems for those institutions that have already created numerous digital objects and wish to participate in the MDL. This issue needs to be addressed in more detail as development progresses.

Several potential MDL participants<sup>30</sup> have acquired excellent hardware and software to support their processes. Their ability to support even small-scale digitization projects for smaller collections may be limited. While education and training will be a central function of the MDL initiative, few potential MDL participants will be able to acquire sufficient hardware or software to produce quality resources. As a consequence, conference participants strongly supported the idea of regional digitization centers similar

---

<sup>28</sup> See Daniel Greenstein (2001) Draft benchmark for digital reproductions of printed books and serial publications (<http://www.diglib.org/standards/draftbmark.htm>; accessed 30 August 2001)

<sup>29</sup> See JSTOR Why images? (<http://www.jstor.org/about/images.html>; accessed 30 August 2001)

<sup>30</sup> For example, the digital initiatives of the University of Minnesota Libraries (Digital Collections Unit at <http://digital.lib.umn.edu/>; accessed 30 August 2001) and the Minnesota Historical Society (<http://www.mnhs.org>; accessed 30 August 2001)

to those established in the Colorado Digitization Project; these regional centers would not compete with or otherwise impede participants from acquiring and managing their own hardware. Regional digitization centers, however, require that selected resources be sent to the center for digitization; this may be difficult for some institutions to consider. Another idea was to develop a mobile digitization lab that could travel to a collection.

Providing collections in a digital environment differs considerably from familiar library or archive stacks. Many potential MDL participants already have hardware to store digital objects, both for archiving high-resolution files and for lower resolution Web presentation; several others have the ability to store and provide access to digital objects. For smaller institutions or those without adequate resources, MDL could establish a digital object repository providing methods to deposit, store, and access the objects. Networked access to shared digital objects that reside on numerous remote servers requires coordination and assurances of consistent and reliable access by participating institutions. This issue will need to be explored more fully in future discussions.

The final and most difficult issue discussed by the standards group was sustainability, the ability to preserve the digital object presentation as standards, hardware, and software develop and change. There are several possible methods<sup>31</sup>, including transfer to a better media,<sup>32</sup> migration of content (e.g., migrations of digital object from one digital format to another as standards evolve or upgrading from one document type definition to another), or emulation of older technology. Undertaking such issues early, following national standards for resolution of the problem, will assure MDL collection stability over time.

Inherent through all the standards discussion, as well as the discussion in other groups, is the need for education and training. Early in the conference, participants discussed the need for a Digital Forum to inform librarians, archivists, historians, educators, and others about digitization and about the MDL. Such a forum would familiarize people with the concepts and terminology and go a long way toward minimizing fears and concerns of potential participants and primary audiences. A forum would also provide a springboard for other training programs for metadata and document type definitions, for digitization processes, for resource enhancement, and for copyright and intellectual property.

### **Audience and Presence**<sup>33</sup>

---

<sup>31</sup> See Caroline R. Arms (2000) Keeping Memory Alive: Practices for preserving digital content at the National Digital Library Program of the Library of Congress from *RLG DigiNews*, 4(3) (<http://www.rlg.org/preserv/diginews/diginews4-3.html>; accessed 30 August 2001); see also Howard Besser (1999) Digital Longevity (<http://www.gseis.ucla.edu/~howard/Papers/sfs-longevity.html>; accessed 30 August 2001) and Alison Bullock (1999) Preservation of Digital Information: Issues and Current Status (<http://www.nlc-bnc.ca/publications/1/p1-259-e.html>; accessed 30 August 2001).

<sup>32</sup> One proposal, adopted by several other digital library projects, was to provide simultaneous conversion to microfilm during digitization.

<sup>33</sup> Participants in the Public Presence discussion: Bill DeJohn, John Butler, Scott Sayre, Joan Roca, Keith Ewing, and Mark Reidell.

Central to the vision of the MDL is establishing a single collective database of metadata describing and providing public access to distributed digital objects that represent the cultural heritage of Minnesota. This single gateway would simplify access and provide entry points for users to all digital resources and functionality of the MDL and participating institutions. Within the MDL, developing a central point of access provides an impetus for expanding participation and further digitization. Externally, the database could facilitate MDL coordination with nationwide developments, for example, with OAI or CMI metadata harvesters.<sup>34</sup>

The public presence discussion group suggested that a short-term solution toward establishing public access would be to create a central Web directory of currently available resources in Minnesota, notably those of the University of Minnesota and the Minnesota Historical Society. Ultimately, however, such a directory would be insufficient for most users.

To meet the variety of need of multiple user communities, MDL must create a database for core and optional metadata and an interface that is capable of customization. At a minimum, the database must be capable of supporting simple and sophisticated user inquiries (including Boolean operations), browsing of selected criteria (e.g., subject, format, geographic location, time period, personal name, etc.), and ability to accept pre-coordinated topical searches (e.g., preset searches based on user preferences). The Colorado Digitization Project's Heritage search<sup>35</sup> demonstrates several of these capabilities.

The functionality of any search depends upon the quality and consistency of metadata available. As MDL develops, it will be important to communicate user expectations and abilities to those charged with maintaining metadata standards. This implies the need to communicate among agencies creating resources using metadata,<sup>36</sup> the need to maintain and analyze search logs for query terminology, and the need to conduct user studies on metadata.<sup>37</sup>

---

<sup>34</sup> OAI, the Open Archives Initiative (<http://www.openarchives.org>; accessed 30 August 2001) develops and promotes technical protocols and standards, collectively called the OAI technical framework, to facilitate access to scholarly research information on the Web. See also John Perkins (2001) A New Way of Making Cultural Information Resources Visible on the Web: Museums and the Open Archives Initiative (<http://www.archimuse.com/mw2001/papers/perkins/perkins.html>; accessed 30 August 2001). CMI, the Research Libraries Group's Cultural Materials Initiative (<http://www.rlg.org/culturalres/>; accessed 30 August 2001), intends to develop and implement the infrastructure to enable integrated, international discovery and use of [digitized cultural] materials.

<sup>35</sup> CDP Heritage utilizes OCLC's SiteSearch as the basic engine; OCLC withdrew support for this application in 2001 and CDP is investigating alternatives.

<sup>36</sup> See, for example, the Minnesota Land Management Information Center's Minnesota Geographic Metadata Guidelines (<http://www.gis.state.mn.us/stds/metadata.htm>; accessed 30 August 2001).

<sup>37</sup> See, for example, the Foundations Project Usability Testing: Dublin Core Metadata and Controlled Vocabulary Study (<http://bridges.state.mn.us/user2study.pdf>; accessed 30 August 2001).

The need to understand and support local and distant users of the MDL and the usability of MDL resources are critical to its success.<sup>38</sup> Every discussion group spoke of the need to conduct user analysis and needs assessment from the beginning and as an ongoing process. A needs assessment should be conducted as soon as feasible to determine user awareness of current digital resources in Minnesota and user expectations from a collaborative project. Part of this assessment would also analyze statistical logs for current projects around the state. An audience analysis,<sup>39</sup> which should include the important native American populations in the state, will provide the MDL with an understanding of current and potential users, their content interests, concerns and applications as well as their access needs and expectations. The research methodologies employed to assess needs, use and usability will guide collection and resource selection, interface design, and strategies for providing user support. The scope of these assessments implies several iterative studies conducted over time as the project develops and matures.

A recently funded joint project of the University of Minnesota, Minnesota State University, Mankato, and Winona State University will explore collaborative responsibility in providing synchronous online reference service<sup>40</sup> to their user communities. This may be the starting point for a broader reference collaborative among academic libraries that could well extend to public libraries and support 24x7 service. If the collaboration is successful, and there is little reason to doubt it will be successful, it may be possible to integrate user support for the MDL or use that experience as a model for a coordinated MDL service. Regardless of how it is accomplished, synchronous online user support should be available to enhance both content and navigation for users. The MDL might need to expand beyond reference librarians to include museum docents, archivists, and other knowledgeable interpreters of resources in order to provide a level of service expected by users.

Enhancement of digital content should be central to the MDL public presence. Inviting scholars and curators to participate by providing context to images<sup>41</sup> can add interest and attract attention. Collection and thematic exhibitions of digital objects, with the ability to include digital objects from multiple collections, could be created to highlight particular resources or in coordination with other events or occasions in Minnesota. Creating exhibitions implies interpretation and re-purposing of digital objects. While it would be possible to allow users to attach individual annotations to any digital object in the collection, conference participants want to present a quality collection, assuring users of the authority (and provenance) and authenticity of digital objects. Curators, archivists, historians, or other knowledgeable people should be responsible for annotations directly

---

<sup>38</sup> See, for example, Dan Greenstein and Denise Troll (2000) Usage, Usability and User Support ([www.diglib.org/use/useframe.htm](http://www.diglib.org/use/useframe.htm); accessed 30 August 2001)

<sup>39</sup> The CDP's Market segments and their information needs (<http://coloradodigital.coalition.org/users.html>; accessed 30 August 2001) provides a useful taxonomy of users and their characteristics.

<sup>40</sup> See Bernie Sloan (1997) Service Perspectives for the Digital Library: Remote Reference Services (<http://www.lis.uiuc.edu/~b-sloan/e-ref.html>; accessed 30 August 2001)

<sup>41</sup> See Minnesota Arts Connected (<http://www.artsconnected.org/artsnetmn/theme.html>; accessed 4 September 2001), a joint project of the Minneapolis Institute of Arts and the Walker Art Center.

integrated into the MDL database for individual objects or collections of objects. However, it would be possible to authenticate k-12 teachers, university faculty, and museum educators to utilize an application that would allow them to create curriculum resources with personalized annotations or commentary connected to the digital objects.<sup>42</sup> This might be particularly effective for linking MDL primary resources (e.g., diaries, letters, and photographs) with information literacy and history components of Minnesota's Profiles of Learning.

The public presence discussion drifted into processes for internal communication and external marketing. Two channels of communication could be pursued in the short-term. First, create an electronic discussion group for all conference participants and other potential participants; this will allow continuing discussion and refinement of issues as future planning sessions emerge. Second, generate interest through local and regional professional meetings (e.g., the Minnesota Library Association conference in October 2001). Marketing the project to the public needs to wait until there is something to demonstrate.

### **Collections**<sup>43</sup>

Minnesota's libraries, archives, historical societies, and museums contain millions of items, among which are numerous unique and valuable resources that collectively represent the cultural heritage and memory of the state and its people. Some of these resources are fragile or subject to damage from even infrequent use; many are inaccessible or unknown to potential users for a variety of reasons. The MDL is conceived as a means of providing broad access through digitization to many of these resources while promoting preservation for the most fragile or the most subject to mutilation. Users might suggest that we digitize everything in our collections, but there are many compelling reasons why this would be unwise.<sup>44</sup> As a place to begin, conference participants recommended that the MDL establish clear selection and collection guidelines.<sup>45</sup>

Providing broad online access to resources requires careful deliberation and consultation with content experts, donors, and other users who may have a stake in the physical item or its content prior to selection and digitizing. Consulting with stakeholders throughout the process, especially with cultural sensitive or donor-restricted materials is good policy.

---

<sup>42</sup> See the Minnesota Arts Connected Arts Collector ([http://www.artsconnected.org/arts\\_collector](http://www.artsconnected.org/arts_collector); accessed 4 September 2001).

<sup>43</sup> Participants in the collections discussion included Nancy Broughton, Sue Ellingwood, Chris Olson, Tom Shaughnessy, Betsy Williams, and Robin Dowden.

<sup>44</sup> See Diane Vogt-O'Connor Selection of Materials for Scanning in *Handbook for Digital Projects: a Management Tool for Preservation and Access* (<http://www.nedcc.org/digital/dman.pdf>; accessed 30 August 2001)

<sup>45</sup> See, for example, the CDP Collection Policy (<http://coloradodigital.coalition.org/select.html>; accessed 30 August 2001).

In addition, establishing intellectual ownership as well as physical ownership<sup>46</sup> or other legal restrictions of an item or collection is a critical step in selecting items for digitization. These issues also should be the topic of an educational workshop for librarians, archivists, and curators whether or not the MDL becomes a reality.

Identifying resources for digitization is a difficult undertaking for a widely dispersed collaborative,<sup>47</sup> complicated by a variety of formats (audio, video, and three-dimensional objects) where digitization standards are yet to stabilize and are often dependent upon proprietary software, and by multiple user audiences for the variety of participating institutions. While the conference participants deferred discussion of selection criteria, several factors arose during the course of the three days. Among these were:

- ?? Value (e.g., rare or unique content, useful to multiple audiences, enhancement of collaborative collections, etc.),
- ?? Demand (e.g., by active multiple audiences, potential for new audiences, etc.),
- ?? Preservation (e.g., reduce handling of fragile materials, reduce risk of theft or mutilation, etc.), and
- ?? Feasibility (e.g., quality of digital object and fidelity with original, ease of access, potential for sustainability, etc.).

Any initial digitization effort by MDL will need to be based on some selection criteria that address these issues.

But where should MDL digitization begin? Images (e.g., photographs, prints, posters, maps, etc.) are fairly easy to digitize if appropriate scanning equipment is available and can generate considerable interest in a short period of time. Researchers will expect more sophisticated interactivity (e.g., the ability to zoom in to see details of maps or prints) with many image files. Some text objects, (e.g., letters, handbills, etc.) can be similarly easy to digitize and, depending on authorship, generate similar interest among users. Both, however, may contain hidden ownership issues that need to be reconciled prior to digitization. Other text objects (e.g., books, pamphlets, playbills, serials, etc.) require more complex digitization processes. Audio objects (e.g., oral histories, traditional storytellers, recorded songs or music, speeches, etc.) and video or film objects (e.g., newsreels, 16mm film, etc.) also require complex digitization processes, but bandwidth issues loom large in the short term for remote access by many potential users. Standards for three-dimensional objects (e.g., pottery, basketry, sculpture, architecture, commercial products, etc.) are still evolving; for the moment bandwidth and proprietary software are limitations.

Another way to answer the question of where to begin digitization is by subject rather than media. Cultural heritage is a difficult concept, involving more than what is uniquely Minnesotan. One possible starting point may be to emphasize Minnesota's history and

---

<sup>46</sup> With conversion of both text and image resources, contracting with a photographer to create an image (digital or chemical) of the original object can result in problems over ownership of the image if contract language is not clear.

<sup>47</sup> See Carl Logoze and David Fielding (1998) *Defining Collections in Distributed Digital Libraries*, D-Lib Magazine, November 1998 (<http://www.dlib.org/dlib/november98/logoze/11logoze.html>; accessed 30 August 2001).

culture before moving into the cultural heritage of Minnesota's ethnic communities. Another possibility would be to work with smaller historical societies and archives to add breadth and depth to already established digital collections at the University of Minnesota or Minnesota Historical Society.

Where to begin digitizing may be best answered through incorporating the previously mentioned user needs assessments into established selection processes used by librarians, archivists, and curators. Parallel to those studies is the need to conduct a survey and identify important, unique, or heavily used resources in collections around the state. Information obtained through these studies will provide a better foundation for determining what content to emphasize in collection-digitizing projects and where to emphasize mini-grant awards for digitizing projects. The discussion group recommended that among the next steps in MDL development there should be a structured collection development process that would guide digitizing projects.

### **Summation**

Several corollary concerns arose during discussions that need to be addressed during the next stage of development.

- ?? Who owns the content of the MDL catalog?
- ?? Who owns digital objects? How will ownership of the original object be noted in the catalog?
- ?? How are rights associated with digital objects managed?
- ?? Will providing access to digital images of material in a collection result in diminished use of the entire collection? Will there be a corresponding loss of revenue if the number of patrons declines?
- ?? Should the MDL assist smaller collections with creating automated catalogs of current collections still dependent upon card catalogs?
- ?? Should MDL compensate smaller libraries for loss of revenue derived through copying?
- ?? How can many libraries cope with two major projects, MnLINK and MDL, occurring at the same time?

These and other policy-related questions can be addressed in the next phase of development.

During a short period between sessions, conference participants, partially prompted by Senator Kelly, explored possible names for the project. Digital library, ambiguous even among information practitioners, fails to convey much meaning to or evoke much response from potential users. Whatever name planners select, it should represent a living project and communicate to users that it is very different from their perception of dusty archives and old libraries. Among the names suggested were:

- ?? Acronyms:
  - MICA: Minnesota Interactive Cultural Archive
  - VENTURA: meaning what?
  - MACH: Minnesota Access to Cultural Heritage

- MONARCH: Minnesota's Online Access to a Rich Cultural Heritage
- ?? Wordmarks:
  - Minnesota Living Cultural Experience
  - Minnesota Interactive Cultural Experience (MICE?)
  - Imagine Minnesota
  - Living Minnesota
  - Minnesota Cultural Heritage Consortium

This report, summarizing the breadth and depth of discussions, demonstrates fulfillment of ambitious conference goals. Conference participants discussed a wide range of issues, including real and potential threats and opportunities. They were able to establish a definition and purpose for the overall project, and delineate the general scope of participation, collections, and audiences. They discussed potential sources of external funding, including private and public grants as well as possible state legislative support. Participants reached a general consensus for a shared policy standard on baseline metadata information. And they identified numerous technical and policy issues that will confront content and access issues as the project develops.

To continue the planning process and prepare a grant for the next stage of planning, conference participants agreed to organize themselves into four working groups. These four working groups were asked to establish some group goals based on this report and discussions held during the conference, to prioritize those goals, to set timelines for deliverables, and to project estimated costs of deliverables. The four groups are:

- ?? Policy and planning: Kristi Tornquist (chair), Mike Kathman, Eileen McCormack, Joan Roca, Christine Clement, and Keith Ewing.
- ?? Standards and training: Charles Thomas (chair), Jason Roy, Eileen Quam, Steve Harsin, Robin Dowden, Joyce Swonger, and Scott Sayre
- ?? Audiences: John Butler (chair), Chris Olson, Steve Harsin, Marian Rengel, Scott Sayre, Sherry Sweetman, and Keith Ewing
- ?? Collections: Daardi Sizemore (chair), Sue Ellingwood, Tom Malefatto, Mark Reidell, Betsy Williams, and Nancy Broughton

It was agreed the chairs of these four groups, with the addition of Keith Ewing, would constitute an executive committee until a more formal structure can be established. The executive committee is charged with writing a second planning grant (deadline for grant completion is 1 December 2001). A follow-up one-day planning meeting for all conference participants is tentatively scheduled for Tuesday, 16 October, place and time to be determined. The executive committee will set the agenda and notify participants. Discussion groups will identify actions and possible deliverables, as well as cost estimates, to be incorporated into the second planning grant.

In general, conference participants, despite reservations about committing their institutions to project participation, were excited by the prospect of sharing their resources through MDL collaboration. The discussions were lively, animated, and provocative. After three long days in meetings, participants were in loud agreement that we should begin something soon and that all potential participants have something to

gain through a collaborative endeavor. At the end participants understood the enormous prospect before them and the tremendous benefits it could yield.

### **Recommendations for Next Steps**

Conference participants agreed to the following next steps:

- ?? Form a steering committee (Status: Executive committee composed of representatives from working groups will work with MAGNOLIA as temporary steering committee.)
- ?? Set up working groups/committees
  - Policy and planning
  - Standards and training
  - Users/audience
  - Collections
  - Rights management (future)
  - Executive
- ?? Working groups will establish goals, prioritize those goals, set timelines for deliverables, and project estimated costs of deliverables (Status: Steering committee met at the University of Minnesota on 5 September and set ambitious agendas for working groups.)
- ?? Write conference report by 14 September 2001 and make available to a wide audience through multiple channels, including on the Web and in print. (Status: Conference report completed on 14 September and distributed to MAGNOLIA members for comment.)
- ?? Create online discussion group/listserv (Status: MDL listserv established at St. Cloud State University.)
- ?? Write second planning grant (absolute last deadline for LSTA funds is 1 December 2001)
  - Grant could fund compilation of a metadata/digitization process and policy manual and people to go out and conduct training sessions.
  - Grant could fund survey of current collections, identify unique or valuable resources for digitization, and promote benefits of project participation.
  - Grant could finalize planning and fund a small proof of concept pilot metadata and digitization project.
  - Grant could fund some combination of the above.
- ?? Develop a plan of actions and realistic deliverables
- ?? Undertake a small test of crosswalk development and loading of metadata and images