Introduction

With a gift from Ameritech, the Library of Congress is sponsoring an open competition to enable public, research, and academic libraries, museums, historical societies and archival institutions (except federal institutions) to create digital collections of primary resource material for distribution on the Internet in a manner that will augment the collections of the National Digital Library Program at the Library of Congress. The National Digital Library is conceived as a distributed collection of converted library materials and digital originals to which many American institutions will contribute. The Library of Congress's contribution to this World Wide Web-based virtual library is called *American Memory* and is created by the Library's National Digital Library Program.

In the 1996-97 competition, applications will be limited to collections of textual and graphic materials that illuminate the period 1850-1920 and that complement and enhance the American Memory collections already mounted in the National Digital Library. The deadline for applications is November 1, 1996 (postmark).

Awards will range from $50,000 to $75,000 for projects that can be accomplished in twelve to eighteen months. Repositories in the United States with collections of primary resource material that are significant for education and research in United States history and culture are encouraged to apply.

The intention of the competition is to encourage applications from a variety of institutions, large and small, public and private. In the final selection among meritorious projects some consideration will be given to the type and size of institution and its geographical location, but the principal criteria will be:

■ The historical significance of the collection's content, as well as its breadth of interest and utility to students and the general public

■ The availability and usability of aids to intellectual access

■ The technical and administrative viability of the project's plan of work

In the first five years of its *American Memory* project the Library of Congress has deliberately experimented with digitizing a variety of materials focused on the political and social history of the United States, with considerable emphasis on images and searchable text. Since standards for the treatment of digital images and text are relatively well-developed, and images have great potential as resources for education and research, applications in the 1996-97 competition will be limited to collections of textual and graphic materials. The guidelines set out for this competition are drawn from the Library's experience with the creation of digital collections.

While the Library of Congress is aware that a number of critical activities are necessary to prepare a collection for digitization and to ensure the quality and future maintenance of digital files, awards from this competition may be used only for the process of digitization (see page 6). The Library has designed the competition to provide an impetus to projects that will enhance the collections of a distributed National Digital Library, while realizing that there will be a substantial contribution on the part of institutions that receive awards. The detailed requirements concerning the procedure for application and the standards to be adopted for the implementation of projects are set forth below.

The National Digital Library Program at the Library of Congress

The historical collections program at the Library of Congress moved from pilot to operational status in 1994, when Internet access to *American Memory* collections was provided for the first time. The purpose of this effort was to enrich the educational process and encourage critical thinking by making available in classrooms, libraries, and homes primary documents of American history. The following is a summary of American Memory collections that illuminate the period 1850-1920.

Now (August 1996) online at http://www.loc.gov/

- *African-American Pamphlets from the Daniel A. P. Murray Collection, 1818-1920* (351 pamphlets offering insight into attitudes and ideas of African-Americans between Reconstruction and the First World War)

- *Selected Civil War Photographs from the Library of Congress, 1861-1865* (1,118 images of Civil War encampments, battlefields, and portraits captured by Mathew Brady and his staff of photographers)

- *The Evolution of the Conservation Movement, 1850-1920* (hundreds of items in a multiformat collection of photographs, printed works, and manuscripts documenting the conservation movement in the United States)

- *Daguerreotypes, 1842-1862* (approximately 600 portrait daguerreotypes, plus a few architectural views and outdoor scenes, by the Mathew Brady studio)

- *Photographs from the Detroit Publishing Company, 1880-1920* (approximately 25,000 postcard photographs of turn-of-the-century America)

- *Early Motion Pictures, 1897-1916* (99 titles from the earliest period of the medium in three groupings: President William McKinley and the Pan-American Exposition, New York City, and San Francisco before and after the great earthquake and fire)

- *The Nation's Forum Sound Recordings, 1918-1920* (59 political speeches by American leaders during World War I and the presidential election of 1920)

- *Walt Whitman Notebooks, 1847-1860s* (four notebooks containing drafts of poetry and prose and personal notes)

- *The National American Woman Suffrage Association Collection, 1848-1921* (a selection of 167 books, pamphlets, and other artifacts documenting the campaign for suffrage for women)

- *World's Transportation Commission Photographs, 1894-1896* (more than 800 images made by American photographer William Henry Jackson documenting foreign countries and their indigenous forms of transportation)

- *Life History Manuscripts from the Folklore Project, WPA Federal Writers' Project, 1936-1940* (approximately 2,900 typescript interviews from the Federal Writers' Folklore Project offering the recollections of Americans from many walks of life)

2. Selected collections planned for scanning, 1996-2000

- *American Broadsides* (tens of thousands of printed advertising pieces, political handbills, theater programs, leaflets, unbound papers and pamphlets, playbills, song sheets, and other non-print material and ephemera, many from the late 19th and early 20th centuries)

- *American Political Prints, 1766-1876* (more than 700 political cartoons, many later than 1850)

- *The American Variety Stage, ca. 1870-1920* (hundreds of materials in many formats illustrating popular entertainment)

- *"California as I Saw It": First-Person Narratives of California's Early Years, 1849-1900* (191 books that tell the story of life and travel in California during the first fifty years of statehood)

- *Folk Music from Northern California, recorded 1938-1940* (more than 800 musical selections performed by European immigrants, together with photographs and other ethnographic documentation recorded for a WPA project)

- *Michigan, Minnesota, and Wisconsin: First-Person Narratives and Local History Reference Sources from the Upper Midwest* (about 100 books that describe community life, occupation, and biography in the late 19th and early 20th centuries)

- *Paper Print Films of the Westinghouse Factory in Action, 1904* (about 20 early films documenting the operation of Westinghouse Air Brake and Electric Motor Company)

- *Panoramic Photographs, ca. 1851-1991* (more than 4,000 large panoramic views of American cities, landscapes, and events)

- *Photographs by Arnold Genthe, ca. 1897-1942* (thousands of images: views of San Francisco's Chinatown before the 1906 earthquake, travel photographs, images of modern dance, and portraits of artists and celebrities)

- *Panoramic Maps, most from 1860-1890* (1,117 birds-eye views of American cities)

The Evaluation Process

The Library of Congress/Ameritech National Digital Library competition project team will review each application and verify that the following objective criteria for eligibility have been met:

- The application comes from a United States institution that is not federal.

- The application has been postmarked by the application deadline (November 1, 1996).

- An original and eleven copies of the application have been submitted.

- All required information has been entered on the cover sheet and budget form.

- The requested award does not exceed $75,000.

- The application conforms to the subject, chronological, and format specifications announced for the 1996-97 competition (United States culture and history, 1850-1920, textual or graphic primary resources).

The Library of Congress will acknowledge the receipt of each application no later than December 1, 1996.

Applications will be evaluated in a three-stage process by independent reviewers external to the Library of Congress. They will first be reviewed for the historical significance of the collection's content, as well as its breadth of interest and utility to students and the general public. The second stage of review will be directed to the bibliographic, technical, and administrative viability of the project's plan of work. Evaluators for these two phases of the competition will be convened by the Division of Preservation and Access of the National Endowment for the Humanities. The third stage of the review will consist of a final panel convened by Dr. Deanna Marcum that will evaluate the most highly recommended projects and make a final selection.

Preparing an Application

The elements of the application should be assembled in the order specified on the verso of the front cover.

1. Application Cover Sheet

Use the form supplied in these instructions. If the cover sheet is prepared using word-processing software, replicate the layout of the form. Do not omit any category of information. The "Brief description of project" should not exceed the space provided (or ten lines, if word-processed.) It may be best to complete the cover sheet last, after all other parts of the application have been prepared.

2. Executive Summary

A one-page précis of the project.

3. Table of Contents

List and supply continuous page numbers for the various parts of the application, including any appendices.

4. The Narrative

The narrative should be concise and written in such a way that it is comprehensible to a general reader. Discussion of technical issues should reflect the understanding of the applicant, and not merely incorporate the language of vendor literature. If an element is lacking in the application, its absence should be explained (e.g., the institution "cannot provide evidence of the use of the collection because it is unprocessed"). The narrative should not exceed ten pages, and should address the following points:

4a. The Collection

- Describe the collection **that is to be digitized in the course of the project.**

- Describe the extent of the collection in commonly accepted measures (e.g., number of titles or items, linear or cubic feet).

- Discuss its significance for both education and research in the history and culture of the United States during the period 1850-1920.

- Give evidence of use of and public interest in the collection in its current format, whether original or a surrogate (such as microfilm or transcription).

- Explain how conversion into electronic format will enhance use of the collection both in the World Wide Web and in other contexts (e.g., scholarly exchange of texts or the use of photographs by publishers).

- Explain the ways in which the collection complements or augments those already mounted in *American Memory*/National Digital Library.

- Describe any other collections already mounted on the Internet that are similar or related to the proposed project.

- Describe the condition of the material and the impact of the process of digitization on the originals.

- Explain any physical treatment that will be necessary to prepare the material for digitization.

- Explain the provisions for care of the collection (and for any film intermediates created in the course of the project) following digitization.

- Identify any privacy or other property rights restrictions attached to the collection and indicate how they have been or will be resolved. (Note: Collections that are digitized as a result of LC/Ameritech awards must be distributable on the Internet.)

4b. Intellectual Access

Digital files cannot be used without appropriate aids to access. Access to documents or images can be facilitated through catalog records or a finding aid, while textual materials may be converted to a form that allows searching for words within the text. An essential aspect of the application is an explanation of the linkage between intellectual access aids and digital images or texts. All applications should include a discussion of the following issues, which have been derived by the Library of Congress from its experience in digitizing its own collections.

- Explain how intellectual access to the content of the digital collection will be provided. Describe the format to be used for catalog records or a finding aid (archival register). For reproductions in the form of searchable texts, discuss whether the information usually contained in a bibliographic record will be provided in structured headers for the text files or in separate catalog records. If some other approach or format is to be employed, explain.

- If catalog records or finding aids currently exist for the material to be digitized, describe them and explain what procedures will be necessary to modify them to support access to the digital reproductions.

- Describe the summary or introductory materials (such as a scope or content note, a bibliography, chronological or biographical information, or a presentation about a particular aspect of the collection) that will be prepared in Hypertext Markup Language (HTML) to provide an intellectual context for the collection.

- Explain the identifier system that will be used to link the catalog record or finding aid to the digital reproductions. Describe the elements in the catalog record or finding aid that refer to the names of the digital reproductions and support links via the World Wide Web to files that contain the corresponding images or texts.

- Explain the level of description to be employed; each link of the description should point to a single digital resource, addressable by a single Uniform Resource Locator (URL) for the World Wide Web. The digital resource may represent a multipart item (such as a manuscript folder or a map reproduced in segments), or a grouping of closely related objects treated as an item (such as a set of photographs of the same subject). If some other approach is to be employed, explain.

- Explain the method and level of subject access. Employ commonly accepted subject thesauri such as LCSH *(Library of Congress Subject Headings)*, AAT *(Art and Architecture Thesaurus)*, or TGM *(Thesaurus for Graphic Materials)*. If some other method of providing subject access is to be used, describe and explain the method.

- Describe how a bibliographic record supporting access to the digital collection will be made available through a national bibliographic utility.

- Describe the experience of staff in providing intellectual access to comparable materials. If project staff are to be hired, describe the required qualifications and indicate the experience of existing staff in supervising such work.

Further discussion of many of these topics may be found in the Technical Notes on Interoperability and Access Aids (pages 11-14).

4c. Digital Conversion Methodology

Describe the overall production process, including production management, estimated throughput, and tracking procedures, as well as the method for carrying out quality review. Production approaches may include the direct capture of original items by means of a scanner or camera, or capture from photographic intermediates, including microfilm, and the conversion of text materials.

- Specify the format or formats for digital reproductions, including the level of resolution and compression for images, and markup schemes for searchable text. Select formats that are in common use and that can be deployed over the World Wide Web. Depending on the class of material, how it will be used, and the need for tonality or color, it may be appropriate to produce images at moderate and/or high resolution. For further discussion of digital formats in use at the Library of Congress and some factors that may influence the choice of format, see the Technical Notes on Formats for Digital Reproductions (pages 8-10).

- If a consultant is to be employed using award funds (non-ARL institution only), discuss the criteria used to select the consultant.

- If an outside vendor is to be used to perform the digital conversion, explain the criteria that will be (or have been) used to select a vendor and monitor the vendor's performance.

- If digital conversion is to be performed in-house, describe the hardware and software to be used (or the criteria to be used in their selection). Discuss the prior experience of staff members in their use.

4d. Provision for Delivery and Network Access

Ongoing World Wide Web access for the general public to the digital reproductions may be provided by the applicant or by the Library of Congress. If the applicant provides access, the collection must be structured so as to permit the Library to integrate the catalog records or finding aid into *American Memory* in a way that will appear seamless to the user.

- If the Library of Congress is to assemble and mount the collection, describe the format and medium for delivery (e.g., write-once CD-ROM) of the catalog records or finding aid, introductory material, and the reproductions. Estimate the disk space required to store the digitized materials.

- If the institution plans to assemble and mount the collection on its own system, explain how it will provide individual addresses for the reproductions that can be resolved via the World Wide Web, using the names imbedded in the catalog records or finding aid; describe the format and medium for delivery of the catalog records or finding aid and introductory materials to the Library of Congress. Describe the current or proposed facilities of the applicant for supporting World Wide Web access and the prior experience of staff in such projects.

- Explain how the information will be maintained and made accessible into the future. Describe the institution's current capabilities to carry out these activities.

- If the institution has previously mounted materials or information on the Internet, provide a Uniform Resource Locator (URL) or instructions for access to this corpus.

4e. Plan of Work

- Provide a plan of work in tabular and/or narrative form that describes the tasks and benchmarks of the project in six-month segments. All project activities must take place within an 18-month period.

- Include all elements of the plan of work necessary to the implementation of the project, not just those to be supported by Library of Congress/Ameritech competition funds.

- Make clear which project staff will be involved in each task and the time commitment of each.

- Discuss how any costs of the project that would not be supported by a Library of Congress/Ameritech award would be met.

5. Budget

Awards will not exceed $75,000. Only costs directly associated with digital conversion may be included in the request (e.g., scanning, modification of cataloging or finding aids to link to digital reproductions, quality control of digital images). Equipment may not be purchased with award funds. Those institutions that do not possess the requisite equipment may use award funds for a vendor to supply scanning services. The institution will be expected to bear costs of administrative overhead, preparation of original materials for digitizing, provision of intellectual access to or preservation of the original materials, and future maintenance of digital materials. Institutions that are not members of the Association of Research Libraries (ARL) may include the costs of a consultant who will assist in the process of digital conversion.

All of the items listed in the budget, whether supported by award funds or cost-sharing contributions, must be reasonable, necessary to accomplish the objectives of the project, auditable, and incurred during the award period. A complete itemization of the projected costs of any third-party contractor should be attached.

Use the budget form supplied in these instructions. If the budget is generated by word-processing or spreadsheet software, follow the layout of the form as closely as possible.

If the applicant is considering using the LC/ Ameritech award to match federal funds from another agency, consult the appropriate official of that agency to determine whether such a match will be permitted.

6. Appendices

Appendices should be attached that contain clearly explained relevant material, such as brief (two-page) résumés describing related experience and qualifications of project staff, job descriptions for project hires, or requests for proposals from vendors. Samples of previous work or of planned access aids may be attached. Indicate the nature of a consultant's responsibilities, if one is used; attach a brief résumé.

Administration of the Award

It will be the understanding governing the award that catalog records and finding aids for new collections will contain certain elements in common with *American Memory* catalog records and finding aids being created at the Library of Congress. Since these practices may continue to evolve, the Library will work closely with awardees to determine specific details (see Technical Notes, pages 8-14).

If funding is awarded, the applicant organization will be responsible for ensuring that the award is administered in accord with the following provisions:

Project activities must commence within the fiscal year in which the award is made (i.e., before October 1, 1997.) Over the period of award activity, an awardee will provide status reports at six-month intervals that describe the accomplishments to date and indicate how successful the awardee has been in meeting the goals described in the project's plan of work. The final cumulative report on the project will be due within three months after the end of the award period.

The award will be made in two payments: 90% of the award will be given at the outset and 10% when the final report is accepted.

Significant changes in the scope, focus, and objectives of the project require prior approval by the Library of Congress. The replacement of the project director or other key staff members who are specifically named in the proposal also requires prior approval by the Library.

Adequate documentation of the time spent by all project personnel on those aspects of the project funded by the award must be maintained by the awardee institution.

When project work will be subcontracted with a third-party vendor, the awardee must develop a contract with the vendor that confirms the costs of the subcontracted activity over the period of the award. The contract should also contain production benchmarks and quality control review procedures.

All procurement transactions should be conducted in a manner that provides, to the maximum extent practical, open and free competition.

Awardees should acknowledge the support of the Library of Congress and Ameritech in all materials that result from or publicize the project's award activities. The Library of Congress will publicize the contributions to the National Digital Library of institutions receiving these awards.

To comply with federal audit regulations, the awardee must have a financial management system that records separately and accurately within its general accounting system the receipt and disbursement of award funds and that monitors all transactions related to the expenditure of these funds against the approved budget. The Library of Congress and its duly authorized representatives will have the right during the awardee's normal business hours to audit and examine books of account and records with respect to the subject matter and terms of the award. The awardee will keep all books, accounts, and records available for at least three years after the completion of the work funded by the award.

Technical Notes on
Formats for Digital Reproductions

Contents:

Introduction

The Library believes that the National Digital Library must support a variety of formats for digital reproductions. Technological options are continually developing. The choice of formats for the capture and presentation of materials depends on the nature of the collections and how they will be used. Several different formats for reproducing text and images are in use for *American Memory* collections released or in production in 1996. They have been chosen to balance quality of reproduction, convenient accessibility for the general public over the World Wide Web, likely longevity of format (using standard formats where possible and proprietary formats only where widely deployed), and production cost. Other institutions may weigh the factors differently, and choose other formats. Other formats (or approaches to dynamic generation of images) that provide effective widespread access to the digital reproductions over the World Wide Web are encouraged. Applicants proposing other formats should be capable of providing access through their own facilities or guiding implementation at the Library.

1. Pictorial materials

For pictorial collections, the Library produces three image types:

Thumbnail
A small GIF image presented with the bibliographic record, to allow users to judge whether they wish to take the time to retrieve a higher quality image.

Reference
The "fetchable" higher quality JPEG image. In current projects, only one reference image is provided; future collections may offer two (or more) at varying levels of resolution.

Archival
An uncompressed TIFF image provided for artifact-free reproduction or for future reprocessing as compression standards change. May be provided to end users as a downloadable file in the future.

Several other organizations have used the Kodak PhotoCD (Image Pac) format for imaging projects. The Library has not had extensive experience with PhotoCD. Applicants wishing to use the PhotoCD technology should either determine how direct access to those images may be provided from their site to WWW clients or plan to reprocess the images to produce GIF and JFIF/JPEG images for mounting at the *American Memory* site.

2. Textual materials reproduced as searchable text and images

Transcribed text can be a tremendous aid to a user seeking a particular word or phrase in textual work. Transcribed text can also facilitate the researcher's navigation of a longer document, especially when encoded with a markup language. The cost of providing perfect or near-perfect transcriptions is very high, however, and, for many researchers, proper understanding of a document may depend upon seeing a facsimile (and in some cases, the original). For these reasons, the Library has experimented with the presentation of manuscript and printed matter items as combined page-images and searchable text. In some pilot collections, separate images of tables and illustrations were provided in addition to or in lieu of page-image sets.

The Library encodes its documents using Standard Generalized Markup Language (SGML), using a document type definition (DTD) based on the guidelines of the Text Encoding Initiative (TEI). Applicants are strongly encouraged to use a TEI-conformant DTD and may choose to use the Library's *American Memory* DTD. Since the Library always places SGML texts online together with bibliographic records or some other access aid, the headers within the documents contain minimal bibliographic information. Full-function SGML viewers for the WWW are not yet available free or as shareware. For this reason, the Library derives an HTML version of the text from the SGML version and places both online. Both copies have active links to page images. The page images included in the searchable presentations employ the formats for tonal and bitonal images described in the next section.

3. Textual materials reproduced as images

The Library has been experimenting with tonal (color and grayscale) reproduction of manuscripts and older printed documents. Original items with a mixture of lighter and darker markings are often more successfully reproduced in a tonal rather than a bitonal image. Typography or line art, however, will often be successfully reproduced in a bitonal image. Bitonal images may provide better printed output. Thus, some collections may warrant the production of both types of images.

3a. Tonal images of manuscripts and printed documents

At this writing, the Library's only online example of tonal document reproduction is the small collection of Walt Whitman's notebooks. As for pictorial materials, the Library produces tonal JPEG images for reference and uncompressed TIFF images as archival versions. A demonstration project to refine a tonal-image approach to manuscripts is under way with a portion of the Federal Theater Project collection. In this latter project, the so-called archival version of the image is tonal and the reference image is bitonal.

3b. Bitonal images of manuscripts and printed documents

The use of the lossless CCITT (Group 4 FAX) compression for bitonal images stored in TIFF format may mean that one image will serve both reference and archiving needs. For some items, however, higher resolution may be desired for the archival version.

3c. The special problem of illustrations in the form of printed halftones

Printed halftones present special problems in reproduction because of interference between the spatial frequency of the halftone dot pattern and the spatial frequency applied by scanning and/or output devices. The interference "waves" caused by the intersection of the two frequencies manifest themselves as aliasing (or moiré pattern) that degrades the image.

There are a number of treatments that can mitigate or correct this degradation but not all are practical in a production-line environment. Possible treatments include the following: descreening and rescreening; cap-

ture at high enough resolution to reproduce the half-tone dots; grayscale reproduction; or "diffuse dithering," a randomization of the dot pattern created by the scanner. Of these treatments, the Library has used only the last, capturing images at 300 dpi and applying the diffuse dithering treatment (at scan time for bitonal images or as a second step for grayscale images). This treatment reduces but does not eliminate moiré patterns. The resulting PCX images print very nicely on a laser printer but do not rescale well for screen display. Archival versions have not been created for images from printed halftones.

3d. Paging through a multi-image set

A number of devices may be employed to present to users the many page images that may be linked to a single reference. The simplest, albeit least elegant, is to create a table of contents or menu (see the current version of the *American Memory* Walt Whitman notebooks collection). Here, an HTML file lists the pages by number and links each one to its reproduction image. A proprietary option that may be appropriate for short documents is to incorporate several images into one file in Adobe's PDF format.

Meanwhile, other organizations have developed systems for their WWW servers that allow paging through a document using sets of images that are displayed "inline" in any WWW browser. If applicants plan to develop such an approach to paging, or to join with the Library as it develops its approach, the project must produce images in the GIF format. This would usually be in addition to higher-resolution images as discussed in the preceding sections.

4. Maps

The Library's Geography and Map Division is developing its approach for digitizing map collections, with the advice of the division's Center for Geographic Information. For the historical maps selected for digitization, the preliminary finding is that good legibility will be afforded for most through tonal images captured at 300 dpi. Archival copies will be stored without compression or with lossless compression. There are many challenges associated with Internet transmission, display, and printing of very large images and the Library has not formulated plans for the presentation of maps in the WWW environment.

Where to find more information

Applicants may wish to consult the WWW site for the Library of Congress/Ameritech National Digital Library Competition (at http://lcweb2.loc.gov/ammem/award/) for further details about formats in use at the Library, including spatial resolution and tonal depth for images. The WWW site also has references to related materials elsewhere.

Technical Notes on Interoperability and Access Aids

Contents:

Introduction

Interoperability will be an essential attribute for easy and effective use of materials in the National Digital Library. While practice within digital libraries is evolving continuously, applicants are expected to propose technical approaches consistent with those in use in August 1996. Alternatively, approaches consistent with standards in advanced stages of development within a national library, archival, or Internet association or community may be proposed by institutions prepared to adapt as those standards take their final form. Elements required for interoperability have been held to a minimum and the suggestions described below (with examples on the competition's WWW site at http://lcweb2.loc.gov/ammem/award/) offer several options.

Each collection must be supported by one or more aids to intellectual access, usually through catalog records, a finding aid (archival register), or the provision of searchable reproductions of textual materials. In the networked, distributed environment envisaged for the National Digital Library, each digital reproduction must be identified in a way that permits retrieval across the World Wide Web. This identifier serves the role played by a "call number" in a traditional library. The identifier supports an active link between access aids and the digital reproductions they describe. The exact linking mechanism will depend on the form of the access aid.

To integrate the collections for which digitization is funded through the LC/Ameritech competition and to facilitate the interoperability of old and new collections, applicants must be prepared to use forms of access aid and linking mechanisms that are compatible with those used by the Library of Congress, usually through consistent use of certain fields or tags. Applicants may, but need not, choose to follow current Library of Congress practices exactly. In the weeks following the announcement of awards, awardee institutions will be expected to work closely with the Library of Congress to ensure that the details of their technical approach allow their materials to be integrated effectively with other *American Memory* collections.

Some suggestions for alternative approaches to identifying items and forms of access aid follow.

Identifiers for digital reproductions

The Uniform Resource **Locators** (URLs) used currently to identify resources on the Internet are unsatisfactory as long-term identifiers, since they usually identify a particular file on a particular computer. When the file must be moved, perhaps when the computer is replaced, the URL will usually change. If URLs are imbedded directly in bibliographic records or finding aids, those access aids must be modified (and the links tested) when the files are moved. Applicants should consider mechanisms that minimize the need for manual editing in the future.

There are currently several proposed schemes for long-term identifiers known as Uniform Resource **Names** (URNs), which would be resolved into physical locators, such as URLs, when users follow links. The Library of Congress offers to help awardees establish URNs for items in their collections using the "handle" service proposed as a URN scheme by the Corporation for National Research Initiatives. Applicants from institutions planning to make use of OCLC's PURL (Persistent URL) system, might consider using PURLs as identifiers. For more information on URLs, URNs, PURLs, and handles, see the competition's WWW site at http://lcweb2.loc.gov/ammem/award/.

Some institutions have taken a local approach to establishing "persistent" URLs that might be feasible for applicants from institutions already using appropriate software. URLs for individual reproductions are actually database queries retrieving a single known item using a unique identifier. The databases being queried might be specialized archives used to manage the digital content or tables mapping identifiers or "logical" names to the URLs representing their "physical" location.

The Library of Congress will be installing a handle-server within the next few months to allow experimentation with URNs. However, *American Memory's* current practice eliminates the need to modify access aids by following rigid rules for naming files and storing them in a directory tree in a way that allows automatic derivation of names for different digital versions of the same bibliographic item. In effect, field 856 contains a "logical" name in fields $d and $f from which "physical" file locators can be derived. If files are moved from one server to another or to a different level of the file hierarchy, the derivation procedure must be changed, but not the individual catalog records. Applicants wishing to follow Library practices and use this approach will be expected to work with the Library closely, after receiving an award, to establish a naming and storage scheme. General information about the storage/naming approach is available on the WWW site, but final details must be worked out after awards are made.

Suggested formats for access aids

1. Bibliographic records in MARC format

All bibliographic records for current *American Memory* collections are in MARC format. Applicants choosing to prepare catalog records in USMARC format need not create full-level records. In addition to the leader and required fixed fields, the MARC fields listed below must be present. No other fields are mandatory for integration with *American Memory,* although the use of author, subject and note fields is strongly recommended to facilitate access.

245 Title statement

856 Electronic access
There is not yet consensus on how best to apply field 856 when several versions of an item are available (such as page-images and searchable text for a document or different qualities of reproduction for an image) or on the subfield to be used for a URN. See the competition's WWW site at http://lcweb2.loc.gov/ammem/award/ for examples of recommended approaches.

985
This LC local field is needed to keep track of the institution supplying the records, a unique code to identify the digital collection, and other administrative details. This field will normally be identical for all records provided by an awardee for a single collection.

659
This LC local field will hold a broad categorization of original material type in a consistent form. For examples, see the competition's WWW site at http://lcweb2.loc.gov/ammem/award/.

At this time, the Library is unable to accept records using the new 774 (Constituent Item) field linked to 856 fields as described in MARBI Proposal 96-4 and integrated into the March 1996 update to the USMARC Bibliographic Format.

2. Bibliographic records following the Dublin Core approach

In March 1995, a "metadata" workshop organized by OCLC and the National Center for Supercomputing Applications was held in Dublin, Ohio. Librarians, archivists, publishers, computer and information scientists, and members of Internet Engineering Task Force working groups approached the task of developing a short list of data elements that could be used by information providers to describe their own resources. The group proposed a set of thirteen elements that is known as the Dublin Core. The elements, designed to support access and retrieval, are: Title, Subject, Author, Publisher, OtherAgent, Date, Identifier, Language, ObjectType, Coverage, Form, Relation, Source. For more detailed information about the Dublin Core proposal, and the followup meeting in April 1996 in Warwick, England, see http://www.dlib.org/dlib/july96/07weibel.html/.

This approach provides an alternative for cataloging that requires less specialized software than the MARC format. However, the Dublin Core concept is still under development and there is not yet an agreed communications format for Dublin Core data. Among the decisions made at the Warwick meeting was that a record format using SGML would be developed for the Dublin Core elements. The Library will be prepared to accept records that use the SGML tagging scheme (document type definition or DTD) that emerges from this consensus-building process. If no firm agreement has been reached on a DTD in the wider group by the time awards are made for this competition, the Library will work with awardees to establish a tagging scheme to be used by the awardees choosing to following this approach. For sample records using the format currently proposed, see the competition's WWW site.

A second suggestion emerging from the Warwick meeting was that the format should be extensible to allow the inclusion of other sets of fields for specific applications. The Library expects to extend the proposed Dublin Core format to support Notes fields (which have been used extensively in the Library's historical collections). Another extension that the Library expects to make, since it may be helpful for some applicants, is a tagging scheme for links to multiple versions of the same item (such as page-images and searchable text for a document or different qualities of reproduction for an image). Applicants should be aware that there is no guarantee that such an extension will be adopted for use at other institutions or in other cooperative projects.

Requirements for interoperability will be similar to those for MARC bibliographic records, with the following fields being mandatory. Details will be worked out in consultation with the Library after awards are made.

Title	must be present
Identifier	required, with details depending on approach proposed for identifying digital reproductions
ObjectType	required to indicate a broad categorization of original material type
Relation	required to support a link to the parent digital collection
OtherAgent	required to identify the institution providing the record

The SGML syntax suggested for a communications format for Dublin Core records is sufficiently simple that special SGML authoring software should not be necessary. It should be feasible to prepare the SGML records as "reports" from many commonly used database software packages. Alternatively, the tags could be added manually using a word-processor or text-editing software.

3. Structured headers in searchable text reproductions

The library has chosen to prepare and manage bibliographic records separately from its searchable text reproductions. However, other institutions may choose to incorporate the bibliographic information usually found in a catalog record into structured headers within the digital texts.

Applicants are strongly encouraged to use SGML document type definitions (DTDs) that conform to the guidelines of the Text Encoding Initiative (TEI) when preparing searchable reproductions of textual materials. For more information on the TEI guidelines, which have been widely adopted for reproductions of historical texts, see the WWW site at http://www.uic.edu:80/orgs/tei/.

A format has been proposed for holding the Dublin Core descriptive elements (see the previous section) in the header section of HTML documents. The proposal (described at http://www.oclc.org:5046 /~weibel/html-

meta.html) may be adopted as a standard, but not necessarily in exactly the form currently proposed. Applicants choosing this approach for cataloging and preparing their digital reproductions should be prepared to modify the headers of the converted texts when the standard syntax is agreed upon.

For interoperability, there will be a few mandatory elements in headers for searchable texts. As for MARC records, a title, a unique identifier for the document, a pointer to the parent digital collection, and an identifier for the contributing institution must all be present. Details will be worked out in consultation with the Library after awards are made. See the competition's WWW site at http://lcweb2.loc.gov/ammem/award/ for sample document headers.

4. Finding aids following the proposed standard for Encoded Archival Description

This proposed standard for structured finding aids marked up in SGML is currently being tested at the Library of Congress and a number of other institutions. The standard will be maintained at the Library of Congress and details are available on the WWW pages of the EAD maintenance agency at http://www.loc.gov/loc/standards/ead/. Because the development of the standard has been supported by the Society of American Archivists and has involved representation from two international bibliographic utilities (OCLC and the Research Libraries Group), it is probable that collaborative online archives of finding aids in the EAD format will be assembled in the future.

To provide links from the finding aid to the digital reproductions, the Library recommends the use of logical identifiers within the main document in conjunction with the SGML "entity" mechanism. Short logical identifiers in the digital archival object element (<dao>) can be resolved into full filenames in an associated entity file. At some future stage, when SGML software recognizes URN identifiers, the entity file could be modified through a global change to use location-independent identifiers rather than physical filenames.

The Library of Congress is currently preparing its

first EAD finding aid that has links to digital reproductions. Sample finding aids describing non-digital collections can be found through the EAD web pages. Because there is currently no free or inexpensive SGML viewer that is fully satisfactory for general use, the Library expects to create HTML versions for browsing, generated automatically from the SGML versions, to support wider accessibility. The Library will be prepared to generate HTML versions of EAD finding aids prepared by awardees.

5. Finding aids marked up in HyperText Markup Language (HTML)

The Library recognizes that some institutions may wish to describe collections in the form of a finding aid document, but prefer to use the simpler HTML standard. The HTML format provides a browsable document that is simpler to prepare and accessible more widely today than a more complex SGML equivalent.

Since HTML primarily facilitates display rather than describing the detailed logical structure of documents, a finding aid marked up in HTML (rather than using the SGML EAD standard) is less capable of supporting refinements in searching such as field qualification. During a search, words in a scope and content note cannot be distinguished from those in a container list.

To minimize the need for later modifications to the finding aid, links to digital reproductions in HTML finding aids should make full and careful use of relative URLs and the <BASE>tag.

Where to find more information

Applicants may wish to consult the WWW site for the Library of Congress/Ameritech National Digital Library Competition (at http://lcweb2.loc.gov/ammem/award/) for illustrative examples and to gain a deeper understanding of the issues relating to the creation of and access to digital reproductions. The WWW site has papers summarizing the Library's experience and current practices as well as references to related materials elsewhere.

Certifications

Information on Certifications Required from Applicants and Awardees
General Requirements

It is the Library of Congress's policy to obtain from all applicants certifications regarding federal debt status and a drug-free workplace. Applicants must also certify that they will comply with the nondiscrimination statutes and implementing regulations.

By signing and submitting an application, the authorizing official of the applicant institution provides the applicable certifications set out below. When a prospective applicant is unable to certify regarding the nondiscrimination statutes and implementing regulations for a drug-free workplace, that institution is not eligible to apply for funding from the Library of Congress. When an applicant is unable to certify regarding federal debt status, an explanation must be attached to the proposal. The explanation of why the certification cannot be submitted will be considered in connection with the evaluators' award determination. Failure to furnish a certification or an explanation shall disqualify the applicant from receiving funding from the Library of Congress.

The certifications are material representations of fact upon which the evaluators will rely when they determine whether to fund the application. If it is later determined that the applicant knowingly provided an erroneous certification or did not comply with the requirements, in addition to other remedies available to the federal government, the Library of Congress may seek judicial enforcement of the certification (nondiscrimination statutes); may terminate the award for cause or default (federal debt status); and may suspend payment or suspend or terminate the award (drug-free workplace).

The applicant shall provide immediate written notice to the director of the National Digital Library if at any time the applicant learns that its certifications were erroneous when submitted or have become erroneous by reason of changed circumstances.

Certification of Nondiscrimination

This certification regarding nondiscrimination obligates the applicant for the period during which the award is given. This certification is binding on the applicant, its successors, transferees, and assignees, and on the authorizing official whose signature appears on the application cover sheet for this proposal.

Certification text: The applicant certifies that it will comply with the following nondiscrimination statutes and their implementing regulations: (a) Title VI of the Civil Rights Act of 1964 (42 U.S.C. 2000d et seq.) which provides that no person in the United States shall, on the ground of race, color, or national origin, be excluded from participation in, be denied the benefits of, or be otherwise subjected to discrimination under any program or activity for which the applicant received federal financial assistance; (b) the Americans with Disabilities Act of 1990 (42 U.S.C. 12101 et seq.), which prohibits discrimination on the basis of handicap in programs and activities receiving federal financial assistance; (c) Title IX of the Education Amendments of 1972, as amended (20 U.S.C. 1681 et seq.) which prohibits discrimination on the basis of sex in education programs and activities receiving federal financial assistance; and (d) the Age Discrimination Act of 1975, as amended (42 U.S.C. 6101 et seq.) which prohibits discrimination on the basis of age in programs and activities receiving federal financial assistance, except that actions which reasonably take age into account as a factor necessary for the normal operation or achievement of any statutory objective of the project or activity shall not violate this statute.

Federal Debt Status

If the applicant institution is unable to certify regarding federal debt status, submit an explanation with the proposal.

Terms used in the federal debt status certification are defined as follows:

Federal Debt: The amount of money or property that has been determined by an appropriate federal agency official to be owed to the United States by any organization or entity.

Delinquent: Represents the failure to pay an obligation or debt by the date specified in a federal agency's initial written notification or applicable contractual agreement, unless other satisfactory payment arrangements have been made by that date, or if at any time thereafter, the debtor fails to satisfy the obligation under a payment agreement with the agency.

Certification text (OMB Circular A-129): The applicant certifies to the best of its knowledge and belief, that it is not delinquent in the repayment of any federal debt.

Drug-Free Workplace

By signing and submitting an application for funding, the applicant institution agrees, among other things, to establish an ongoing drug-free awareness program; to publish a statement notifying employees that the unlawful manufacture, distribution, dispensing, possession, or use of a controlled substance is prohibited in the awardee's workplace; and to give this statement to each employee to be engaged in the performance of the award. The program and policy statement must be in place within thirty days of the date the award is issued.

As an awardee, the applicant institution will be considered in violation of the drug-free workplace requirements if it falsely certifies, fails to carry out the requirements of the certification, or fails to make a good-faith effort to maintain a drug-free workplace.

In the application, the applicant must either identify the place(s) where the award activities will be carried out or must keep this information on file in its office so that it is available for federal inspection. Workplace identification shall include the actual address of buildings (or parts of buildings) or other sites where work under the award will take place. Failure to identify all known workplaces constitutes a violation of the awardee's drug-free workplace requirements.

The terms used in the drug-free workplace certification are defined as follows:

Controlled Substance: A controlled substance in schedules I through V of the Controlled Substance Act (21 U.S.C. 812), and as further defined by regulation at 21 CFR 1308.11-1308.15.

Drug-free Workplace: A site for the performance of work done in connection with a specific award at which employees of the awardee are prohibited from engaging in the unlawful manufacture, distribution, dispensing, possession, or use of a controlled substance.

Employee: The employee of an awardee directly engaged in the performance of work under the award and all temporary personnel and consultants who are directly engaged in the performance of work under the award and who are on the awardee's payroll.

Applicant: An institution that applies for or receives an award directly from the Library of Congress under the LC/Ameritech National Digital Library competition.

Certification text (Drug-Free Workplace Act of 1988):

(A) The applicant certifies that, if given an award under this program, it will continue to provide a drug-free workplace by

 (a) publishing a statement notifying employees that the unlawful manufacture, distribution, dispensing, possession or use of a controlled substance is prohibited in the awardee's workplace and specifying the actions that will be taken against employees for violation of such prohibitions;

(b) establishing an ongoing drug-free awareness program to inform employees about (1) the dangers of drug abuse in the workplace; (2) the awardee's policy of maintaining a drug-free workplace; (3) any available drug counseling, rehabilitation, and employee assistance programs; and (4) the penalties that may be imposed on employees for drug abuse violations occurring in the workplace;

(c) making it a requirement that each employee to be engaged in the performance of the project be given a copy of the statement required by paragraph (a);

(d) notifying the employee in the statement required by paragraph (a) that, as a condition of employment under the award, the employee will (1) abide by the terms of the statement; and (2) notify the employer in writing of his or her conviction for a violation of a criminal drug statute occurring in the workplace not later than five calendar days after such conviction;

(e) notifying the Library of Congress in writing within ten calendar days after receiving notice under subparagraph (d)(2) from an employee or otherwise receiving actual notice of such conviction.

(f) taking one of the following actions, within 30 calendar days of receiving notice under subparagraph (d)(2), with respect to any employee who is so convicted: (1) taking appropriate personnel action against such an employee, up to and including termination consistent with the requirements of the Rehabilitation Act of 1973, as amended; or (2) requiring such employee to participate satisfactorily in a drug abuse assistance or rehabilitation program approved for such purposes by a federal, state, or local health, law enforcement, or other appropriate agency;

(g) making a good faith effort to continue to maintain a drug-free workplace through implementation of paragraphs (a), (b), (c), (d), (e), and (f).

B) The applicant shall either identify the site(s) for the performance of work done in connection with the project in the application material or shall keep this information on file in its office so that it is available for federal inspection. The street address, city, county, state, and zip code should be provided whenever possible.

Library of Congress/Ameritech National Digital Library Competition
Application Cover Sheet

Institution or Organization

Name _____

Address _____

_____

_____

(city)

Project Director

Name _____

Tel. no. _____

E-mail address _____

(state)   *(zip code)*

Descriptive title of project _____

Brief description of project (do not exceed space provided)

Authorizing official

Name_____

Title_____

Address _____
    (street address)   *(city)*

Tel. no. _____

E-mail address _____

(state)   *(zip code)*

Employer identification number _____

Member of the Association of Research Libraries ☐ Yes ☐ No

Certification:

By signing and submitting this application, the authorizing official of the applicant institution is providing the applicable certifications regarding the nondiscrimination statutes and implementing regulations, federal debt status, and a drug-free workplace, as set forth on pages 15-16 of these guidelines.

_____

(printed name)

_____

(signature)   *(date)*

NOTE: Federal law provides criminal penalties of up to $10,000 or imprisonment of up to five years, or both, for knowingly providing false information to an agency of the U.S. government.

Library of Congress/Ameritech National Digital Library Competition
Budget Form

Applicant Institution _____

Project Director _____

Requested Grant Period

From _____ to _____
        mo/yr   *mo/yr*

The three-column budget has been developed for the convenience of those applicants who wish to identify the project costs that will be charged to LC/Ameritech funds and those that will be cost shared. FOR LC PURPOSES, THE ONLY COLUMN THAT NEEDS TO BE COMPLETED IS COLUMN A. The method of cost computation should clearly indicate how the charge for each budget item was determined. If more space is needed for any budget category, please follow the budget format on a separate sheet of paper.

1. Salaries and Wages

Provide the names and titles of principal project personnel.

name/title or position	method of cost computation	LC/Ameritech Funds (a)	Cost Sharing (b)	Total (c)
_____	_____	$_____	$_____	$_____
_____	_____	_____	_____	_____
_____	_____	_____	_____	_____
_____	_____	_____	_____	_____
_____	_____	_____	_____	_____
_____	_____	_____	_____	_____
	Subtotal	$_____	$_____	$_____

---

2. Fringe Benefits

If more than one rate is used, list each rate and salary base

rate	salary base	(a)	(b)	(c)
_____ %	of $_____	$_____	$_____	$_____
_____ %	of $_____	_____	_____	_____
	Subtotal	$_____	$_____	$_____

3. Services

Include the cost of duplication and printing, long distance telephone, postage, consultant fees, and other services related to project objectives. Provide an itemization of subcontract costs on this form or on an attachment.

item	basis/method of cost computation	(a)	(b)	(c)
		$_____	$_____	$_____
		_____	_____	_____
		_____	_____	_____
		_____	_____	_____
		_____	_____	_____
	Subtotal	$_____	$_____	$_____

4. Other Costs

Include supplies and materials, equipment purchases, and other items not previously listed. Please note that "miscellaneous" and "contingency" are not acceptable budget categories.a

item	basis/method of cost computation	(a)	(b)	(c)
		$_____	$_____	$_____
		_____	_____	_____
		_____	_____	_____
		_____	_____	_____
		_____	_____	_____
		_____	_____	_____
	Subtotal	$_____	$_____	$_____

5. Total Project Costs

$_____ $_____ $_____

Table of Contents

Application deadline
November 1, 1996 (postmark)

Award announcement
March 31, 1997

Checklist of Application Elements

An application should contain the following elements arranged in this sequence:

☐ Application Cover Sheet

☐ Executive Summary

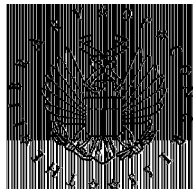☐ Table of Contents

☐ Narrative

☐ Budget

☐ Appendices

Applicants should submit twelve copies of the completed application. One copy should have an original ink signature. All pages of the application should be printed single-sided, single-spaced, and numbered consecutively. Applications should be stapled or clipped, but not enclosed in binders. The narrative of the application should not exceed ten pages. No extensions to the deadline will be granted.

The Library of Congress
Washington DC 20540-4860
Telephone: (202) 707-1087
Fax: (202) 707-3566
E-mail: bpau@loc.gov
http://lcweb2.loc.gov/ammem/award/

National Digital Library Competition

1996 Guidelines and Application Instructions

Ameritech