

Werk

Titel: The Library of Congress

Ort: Graz

Jahr: 1995

PURL: https://resolver.sub.uni-goettingen.de/purl?514854804_0005|log37

Kontakt/Contact

[Digizeitschriften e.V.](#)
SUB Göttingen
Platz der Göttinger Sieben 1
37073 Göttingen

✉ info@digizeitschriften.de

The Library of Congress

Technological Strategies for a National Digital Library

Summary of Conference Proceedings

September 1-2, 1994

Executive Summary

In the world of ubiquitous electronic information, digital libraries will link Americans to vital information. To consider the role of the Library of Congress in the creation of a National Digital Library, a two-day conference with selected industry and academic leaders in technology was held September 1-2, 1994.

Conference participants stressed that many of the technical questions concerning digital information will not be resolved quickly. The Library, however, must begin to digitize its collections. It should begin by digitizing selected items from its Special Collections. Such an effort is an investment in the country's future educational and technological development. Once digital collections are accessible electronically, researchers will use them, creating an even greater demand for digital knowledge. As use grows, information specialists will find new and better ways to provide access to digital collections. Eventually, all the various experiments and research will merge into a system that makes it easy for users to access and use Library of Congress collections electronically. The Library cannot wait until all questions are answered before beginning to contribute to this effort. The Library, therefore, should move ahead on projects that contribute to the foundation of the National Digital Library.

The Library of Congress must play a leadership role. The Library should push the nation's current library systems toward the future. By digitizing items at the highest level of quality possible, the Library positions itself to take advantage of future technologies. The Library should play a key role in the development of standards for digital documents. Investments in digitization will provide greater returns if goals for a digital library are shared.

Collaboration is vital if a National Digital Library is to flourish. Industry experts can help identify the best technological solutions through research. They

can aid the Library in the dissemination of its collections to educational institutions across the country. Companies will benefit from the Library's digital offerings by creating products that add value to the digital items provided by the Library, creating commercially viable products.

As the digital world grows, complex questions of intellectual property rights will continue to present difficulties. The Library's Copyright Office must play a crucial leadership role in discussions and resolution of copyright issues.

The rich collections of the Library of Congress are a national treasure. By putting the most useful of them online, the Library can help share these riches with people across the nation and around the world.

TECHNOLOGICAL STRATEGIES FOR A NATIONAL DIGITAL LIBRARY

Summary of Proceedings

Introduction

On September 1, 1994, the Library of Congress convened a two-day meeting to identify technological strategies for building a National Digital Library and to define what the Library's role in those efforts might be.

The Librarian of Congress, James H. Billington, presided over the conference.

Participants included leaders in the field of digital technology from industry, non-profit, and academic organizations. Library staff who participated have major responsibilities for planning and implementing the Library's efforts to develop collections for a National Digital Library.

Finding better ways to manage electronic information is a challenge that calls for collaboration among knowledgeable corporate and government information scientists. Participants offered specific suggestions for potential collaborative efforts and spending large amounts of money on these experiments. The conference participants dedicated to the conservation of the Library's holdings, the nation would move closer to having a core database for a National Digital Library.

Prepared Remarks

JAMES H. BILLINGTON
Librarian of Congress

Dr. Billington highlighted the benefits of a National Digital Library. His objective is to make the significant and unique collections at the Library of Congress available electronically to a broader audience than those who can travel to Washington. Of special interest are those items that further the study of American history and culture and an understanding of the nation's democratic process.

The Library does not intend to usurp the role of the private sector in delivering digital data. If private industry can provide a service, the Library will work with those companies. However, if important services that are in the national interest are not provided by industry, the Library will strive to fill that void.

The Library does intend to act as a consensus-building force, particularly in the area of intellectual property rights. The Clinton Administration's Information Infrastructure Task Force is considering the issues of intellectual property rights in the electronic world. The information superhighway needs "traffic rules." Intellectual property rights will help provide these.

The Library has begun to digitize its collections, and plans, in collaboration with other repositories, to make more than 5 million items available electronically by the year 2000. As the Library makes its materials available digitally, retrieval will present a significant challenge. The Library will explore ways to improve finding aids for digital information to enhance the retrieval of information.

Dr. Billington expressed hope that the information superhighway would reinforce the best values and productive dynamism of American society: "This emerging interactive world does, after all, engage the active mind in intellectual-calisthenics and in creative interplay with useful information." But the cargo transported on the information superhighway will greatly affect how beneficial the new infrastructure will be to society. "If the highway provides just entertainment and high-priced information on demand, the gap will probably widen between the information haves and have-nots," Dr. Billington emphasized. Americans must not forfeit the inexpensive and broad access to knowledge that public libraries and public education systems offer today, and, with the help of the new technology, can offer tomorrow.

The Library of Congress expects to routinely receive significant amounts of new material in digital form. Items such as films, music, encyclopedias, legal records, maps, scientific paper, and government documents will be stored in electronic collections. The Library also will work to digitize its most useful existing paper, audio, and film collections for electronic distribution. Dr. Billington is particularly interested in providing these electronic collections to schools and local libraries without charge or at reduced rates: "This effort has great potential to help schools achieve the high national education standards newly approved by Congress, and can help the country, we think, to come out of the educational slump that began during the 1960s."

The Library of Congress already has become a major presence on the Internet. More than 7,000 network visits are made to the Library each day. The Internet now offers access to more than 40 million Library records, including the entire Library of Congress catalog, summaries and status of federal legislation, copyright registration records, and abstracts and citations for foreign laws. The images and accompanying texts from the Russian Archives, Columbus Quincentenary, Treasures of the Vatican Library, Scrolls from the Dead Sea, and the African-American Mosaic. More than 400,000 electronic visits to the Russian exhibit alone have been counted.

The transition to a National Digital Library already has begun. A five-year pilot of the Library's American Memory project is nearing its conclusion. That project has made possible the digitization of 24 Americana collections, including prints and photographs, manuscripts, sound recordings, and motion pictures. As part of the test, 44 schools and library sites received the multimedia collections. By the year 2000, the Library in collaboration with other institutions plans to offer 5000 to 1,000 collections containing more than 5 million Americana items.

The National Digital Library is a means to provide unenhanced archival transfer of the Library's collections. Private sector entities can add value to the collections, and local groups or individuals can reshape them in meaningful ways that provide educational value. The electronic services will expose library users to new technology and old values, memory of the past and imagination for the future.

Several projects already under way are moving the Library of Congress toward a National Digital Library. They are:

- The Electronic Copyright Management System, a testbed supported by the Library and the Advanced Research Projects Agency. The project involves developing and evaluating a system for electronic copyright deposit, registration and recordation.
- The Electronic Cataloging in Publication project, which is testing online transmission of galley proofs from several publishers using the Internet. The test is designed to ease the preparation of cataloging

information and build a foundation for an electronic library of machine-readable books.

In January 1995, the Library added a major electronic source of congressional information called THOMAS; in honor of Thomas Jefferson.

CONFERENCE SESSIONS

The Bernstein Archives: The Working Collections Model for the Conference

DEANNA MARCUM / JAMES PRUETT

Director of Public Service and Collection Management I / Chief of Music Division

Digitization

The Leonard Bernstein Archives will be the first part of a digital archives of American music that the Library is creating. The digital music archives, which contain 25,000 linear feet of historical materials donated by the Bernstein family, will be available to other libraries, schools, and cultural organizations. The Library has drawn an advisory council from six other institutions with significant music collections help guide the transition of the collection to digital form. Council members are working to identify collections that will complement the Bernstein materials, and to determine the best way to digitize those materials and make them available to the public.

The digitization of the Bernstein Archives is one step in the Library's progression toward providing electronic access to its special collections. The special collections comprise 75 million of the 10 million items in the Library's collections. The public can discover many of the items in the collections only through reference librarians or specialists in those particular areas, because only a small percentage of the special collection items are under bibliographic control.

Conversion

The Bernstein Archives provide a representative sampling of the range of materials to be converted from paper to electronic form, although they are far less paper-based than other holdings in the Music Division. Participants studied samples of the collection, which includes 117 scrapbooks kept throughout Bernstein's career, date books, personal letters to family members and colleagues, pencil drafts of musical compositions, concert programs, and record

jackets, among many other special-format items. The collection presents a number of challenges to digitization because of the great variety of materials within it.

In studying the collection materials, participants questioned the usefulness of providing such material digitally without the contextual information that makes the data meaningful.

Copyright

A number of questions arose regarding restriction and origination of digitized materials:

- Once the sound recordings or the sheet music are digitized and made available on the Internet, can anybody then use the music?
- Can someone essentially use it as if it were in the public domain?
- When works are created digitally, what constitutes the original version of the work?
- How will people distinguish between the original and the copy for preservation purposes?
- The Library of Congress has a statutory responsibility for enforcing the copyright laws and requiring the deposit of copyrighted material. But it also has a seemingly contradictory role, which is to promote the arts and sciences by making the material available to the public as fully as possible. The Library must find a balance between access and protection.

Session I: Digitizing Multimedia Collections

LAURA CAMPBELL

Director of Library Distribution Services, Moderator

Digitization of the special collections in the Library of Congress will provide access to people who cannot travel to Washington to view the materials. About 80 percent of the special collections are housed in Library buildings in Washington, and many of the collections are unique. By offering electronic access, the Library can improve the educational and research benefits derived from these works.

During the last five years the Library has been conducting a pilot project to test methods for digitizing its special collections. Under the pilot, the Library has converted 220,000 items, including Mathew Brady Civil War photographs, early films of San Francisco before and after the great earthquake, political cartoons, and documents from the Continental Congress and Constitutional Convention.

During the tests, certain conversion challenges became clear. It is difficult to convert rare documents, because many cannot be sent through a document feeder or a book box for filming. Examples of other troublesome materials include 78 rpm records, oversized maps, and early video images. Creative solutions require time to develop, and the digitization process will not move quickly for materials such as these.

As the Library converts a large body of historical materials, some of the challenges that must be resolved include:

- Achieving a high level of productivity
- Preserving fragile artifacts
- Handling a variety of original formats
- Retaining the original
- Providing security for the original
- Verifying the quality of what is converted
- Keeping a preservation copy
- Adding finding aids and tagging to the material
- Adopting standards and developing a common method for conversion

The first goal of digitization is access, but the ways in which the public uses the information will affect the methods of storage and the finding aids that are needed. In a project at Cornell University, the staff compared usage of digital images versus digital text data from chemistry journals. Users most often chose to browse through the page images rather than the text keyed from the same, original page. This leads to questions about how library collections should be presented in the digital world. Should a page image be preserved or should access to the content itself be the primary goal. The answer is probably both, depending on the distinctive qualities of each item that is scanned into a computer system.

Access

There is great tension between providing electronic access to digital library collections and at the same time preserving them. The Library of Congress must find a balance. One participant suggested that today's generation will judge libraries by access and future generations will judge them by what they collected. The Library's goal is to provide more access to its collections while preserving them for future generations.

Costs

Digitization is costly; the Library cannot depend solely on its budget allocations to pay for the digitization of its collections; collaboration will be the rule. One participant suggested reducing library building costs by sharing the costs of digitization. Two different university libraries recently built new book stacks for about \$20 and \$30 per book. If both libraries owned the same book and could find a way to scan it legally and to share it, they would save storage space and cut down on future buildings costs. Such a strategy would work, however, only for books that could be destroyed because they had no artifactual value.

Some university libraries have projects under way to scan decades-old academic journals. Such materials lend themselves to digitization and destruction of the original paper copies.

Other libraries are working on developing "smart" optical-character recognition (OCR) or structured, editable, readable documents from scanned works. The Library of Congress has more than 300,000 reels of microfilm, including many presidential papers. Finding an efficient way - such as smart OCR - to convert that material to readable, searchable electronic text is of particular importance.

Another cost issue is determining whether the Library wants to spend money to clean up digital works (e.g., the fuzzy recordings offered by early sound technology). One participant suggested allowing individual users or customer organizations to clean up a work and send back the modified copy for the Library's collections. Audio tapes and artwork, for example, often have deteriorated and could benefit from restoration.

Local users also could play a role in creating specialized packages of Library information. Some are already developing very specialized CD-ROM multimedia packages in their homes. If the Library nurtured such efforts, this cottage industry could help the Library get its special collections out to the public.

Another way to manage digitization costs is to provide digitization-on-demand, much as organizations now provide print-on demand. Rather than digitize an entire collection, the Library could digitize items as requests are made. Such a strategy leads to other questions, however, such as who will play for the digitization and whether it could be performed quickly enough to meet the requester's needs. One participant suggested having a service that would require a requester to pay a nominal processing fee of say \$25 for the digitization of a book on demand. A digital copy of the book also would go into the Library's collections. The costs of digitization are subsidized and the Library increases its digital collections based on usage. Obstacles to this approach include the high cost of digitizing little-used works. If the demand is low, processing fees are unlikely to equal actual digitization costs.

A similar approach would be to study usage patterns and digitize the most-used items. This approach raises questions of access, however. Would the usage of seldom used items increase if they were available over the Internet? The answer may be "yes." If so, basing digitization priorities on usage figures is misdirected.

The Library, therefore, needs several methods that work together. The first part of an adaptive, self-correcting system is scanning requested items. The second part is scanning materials that are deteriorating and have been determined to be worth preserving. Finally, a group of experts should create a list of items in priority order for digitization. Based on the available budget, those items would then be scanned.

The consensus was that the Library needs to fund some digitization and seek alternatives to subsidize the remaining digitization process.

Cataloging

Cataloging is a major component of digitization costs. Should graphic or sound images be cataloged at the item level? If so, how can millions of items be cataloged so that the cataloging is useful, cheap, and quick? If these formats are not cataloged, then how are they to be indexed?

Some participants' suggestions include:

- Making the graphic images available and letting users add their own ancillary data that could be used as low-level finding aids and would be distinguished in some way from those officially created by staff.
- Collaboration with telephone companies could further the goal of audiotape cataloging. AT&T's labs, for example, are working on speech recognition projects for cataloging voice mail messages so users can retrieve certain messages automatically. Such technology has the potential to be used for audiotape cataloging.
- Cooperation among government agencies that are working to develop cataloging techniques for digital libraries.

Searching

Digitizing information is an important step forward, but how will people find the information in the growing stream of data on the Internet? The digital collection of the Library of Congress must be usable and not just available. Searching tools are of vital importance in bringing these collections to the public.

As librarians and information scientists debate the best methods for searching electronic collections, it must be kept in mind that the digital document is a new entity, which characteristics of its own, including its evolution, usage pattern, interface, and demand. Searching tools must be flexible enough to change as the uses of documents change.

Session II: Organizing Multimedia Collections

SARAH THOMAS

Director of Cataloging, Moderator

The Library of Congress's cataloging procedures are based on Charles Cutter's *Rules for a Printed Dictionary Catalog* (1904). Those rules of organization may or may not be applicable in the modern, digital library environment.

The original purpose of a catalog was to enable a person to find a book by author, title, or subject, to show what a library had, and to empower a person to choose a book based on this information. The catalog worked for books and serials, but was less effective for other materials. Those other materials now can be incorporated into the digital environment. Technology can help the Library find new strategies to access such materials.

Challenges in Cataloging

As libraries organize their multimedia collections, a number of questions must be addressed:

- What technologies are available to identify, describe, and analyze digitized collections of multimedia materials?
- How can libraries reduce the human resources investment in collection analysis and in handling of materials?
- What is the value that the library expert can add to the organization of this material?
- What should the goals of technology be, and what should librarians and experts be contributing?
- Should the Library take published indexes to material in its collections and build from them?
- How can technology help the Library transcend or encompass the various perspectives of specialists and generalists who will consult the collections?

- What advances are there in the use of non-standard character sets that can increase access to materials in non-western languages?
- What organizations have similar problems?

Participants discussed the organization of collections of material such as photographs. The Library of Congress traditionally provides general access to photographs.

Few interpretations are provided. Users, therefore, must decide how photographs are to be interpreted and used.

Photographic Resources: Strategies for Searching and Public Input

The Library must determine how much involvement is needed to describe and index photographs. Photo files could be searched by criteria such as the printing medium or the subject of the photo, but most searching of photographs is best done by browsing. One participant pointed out that photo searching often involves nonverbal criteria. Photo database services provide such searching for a fee. If the Library becomes more involved in photo indexing, it might be seen as competing with this industry.

The Library of Congress generally does not try to interpret photos, but rather provides the images and lets users make their own interpretations. When the Library first created a video disk for photographs, it included 25,000 negatives from the Detroit Publishing Company with no cataloging data. Users browsed through the images, but then they wanted to know what was in the image. That created a problem. In many cases, the subject matter of photos and the categories of subjects can be identified only by the human intervention of catalogers and subject specialists. Providing the images without cataloging data or captions might not meet the needs of Library users.

The University of California's Computer Science Department is working on alternative strategies for searching massive textual databases using loosely structured queries. Use of such a strategy for searching photo collections is dependent on someone writing brief descriptions of the photographs, but it could provide an alternative indexing option.

If more indexing is needed, there are several ways to get the public more involved in the digitization process. Photos are often scanned and uploaded by individual users to the Internet. Users of these photos then upload their own descriptions of the photos, providing rudimentary material for potential searching. One suggestion is that once the Library disseminates its files, researchers could add their own comments to the photo database. Some comments would be valuable; others are likely cataloging, but could provide an alternate indexing strategy. The consensus among conference participants was

that if local input was allowed, the system should have two regions for data storage: a region with images and descriptions approved by the Library of Congress; and a region in which users are on their own and understand that the data might not be officially sanctioned.

Finding Aids

As digital collections grow, libraries will seek creative ways to use finding aids to add value to those electronic holdings. Searches are useful only if users can find the material they are seeking. Therefore, new ways of cataloging must be explored. Many of the items in the Library of Congress's special collections do not lend themselves to traditional cataloging but are best described in a few sentences with the description then searched as full-text.

As an electronic system grows, the search criteria and the success rates tend to diminish, according to one participant. The Library, therefore, must be careful about adding excessive descriptive information to be used for searches. The consensus among the participants was that if the Library of Congress could make the raw digital data available, research into new searching technologies would follow.

Session III: Navigating Multimedia Collections

HERBERT S. BECKER

Director of Information Technology Services, Moderator

As the Library of Congress moves to make massive amounts of multimedia items available over networks, it must consider the challenge these new digital collections present for navigation. Sound and image databases are particularly challenging to searchers. For example, how can a digital record be searched if it has no finding aid, structured text, or unstructured description of the record? What kind of global searching mechanism do libraries want and what kind of assumptions would they make about the dynamics that occur a search session?

Discussion during this session centered on what Library visitors are seeking. What are the expectations of the Library's digital collections audience? Studies have shown that 95 percent of the people using public libraries are browsing. At college and university libraries, 25 percent of the people are searching for known items. At special libraries, 50 percent are doing known-item searches. Because known-item searches tend to be focused, the challenge is to define the subject of an item. Users come into the Library with many preconceived notions about what they need. As their backgrounds vary, so will their definitions of the same subject and therefore the words used to search for items in a given category.

It is difficult to distinguish who the Library's users are. It is easier to define what the Library's users use. The book collections are most frequently used items; users primarily want English-language material published in the last five years. This usage pattern does not necessarily indicate the most important use of the Library's collections, however. Many quick visits account for the frequently used items, but the serious scholar mines the depths of the Library. These scholars often stay for months or years because the Library's collections are vital to their research. The Library must find a way to meet the needs of both kinds of visitors - the quick book search and the requests of the long-term scholar - and provide them with searching aids that support each researcher's needs.

One option is for the Library to support free text and structured searching, including dynamic relevancy ranking. Also, automatic clustering and dynamic classification as aids to searching should be tried. Those strategies will support people who are looking for a particular item and want to do searches that will point them to that item. For people trying to broaden their cultural understanding, interfaces that enable them to browse through the data and

choose items will be helpful. The intellectual process by which people use analogical thinking to synthesize data into knowledge must not be forgotten. People move from one subject to another when thinking or communicating. Similarly, electronic searching should be able to model this intellectual process.

Visual Navigation

As graphical user interfaces continue to overpower the text-only world, two-dimensional and three-dimensional viewing technologies are growing. More and more computer systems are based on point-and-click technology that lets the user navigate information visually. One participant brought up the concept of a "library in a cave". The virtual reality community has been developing the idea of a cave, where the user views presentations displayed in a three-dimensional environment. The question is: what should that experience include? Some of the features should focus on browsing, others on searching. The options for library navigation in an electronic world could go way beyond the mundane experience of viewing plain text on a computer terminal. Such creative interfaces expand the options of how students learn and how researchers find information. Libraries must consider experimenting in this area.

References to resources

As users navigate digital multimedia collections, the Library of Congress will not be the sole source of information. The Library must define a niche for itself as one of many sources of electronic information. Participants had two different visions of the Library's future role as an information guide. Some said the Library of Congress should be the first place people look for comprehensive answers. In such a situation, other libraries would feed information about their networked resources to the Library of Congress, which would act as an online directory of other libraries' holdings. Other participants felt the Library should not serve as a directory. Instead, they believed the Library of Congress should be the place users seek information after educating themselves about the options at the local level. Referral is going to be important, because a database with 10 million records is poorly set up to answer basic questions such as, "I'd like a book about Russian history", or "I need an introduction to physics". As a cost-saving measure, the Library could send users to their local public libraries first, and let the local libraries then route them to the Library of Congress when necessary, just as it does now with written reference requests.

Technology levels

If the Library's larger goal is increased access to its collections, then it must focus on the technology level of its users and, consequently, select appropriate interfaces for its data to suit those levels. Ideally, the interface should open access to users at many technology levels. Users on the Internet probably cover four or five different technology levels; these must be taken into account. The library will have to make some tradeoffs between an interface that requires high-end hardware and connectivity, and an interface designed for an ASCII terminal and lower speeds. Sophisticated users are going to campaign for the interface that delivers the most without regard to hardware constraints, but the Library must be careful not to leave the majority of its clients behind.

The technology problem can also become a chicken-and-egg issue, one participant noted. If the Library becomes too preoccupied with the common denominator, it will not be in a position to push the state of library technology forward. Another participant added that the Library should digitize items at the highest resolution and disseminate its digital offerings using the best state-of-the-art technology possible within cost constraints. The Library could push industry to create better interface technology. The library community also must push users to get better hardware, software, and connectivity. If the Library could move slightly ahead of the common denominator, for example by using multimedia workstations with sound capability, public libraries and schools would be likely to follow and to acquire the compatible, new technology. The tangible payoff to schools and public libraries would be access to multimedia programs put out by the Library of Congress. The key is for the Library to store its data in the highest possible resolution, in a format that lends itself to reformatting and redistribution later. Aiming high does not guarantee that the resolution will be high enough. The best computer technology can become outdated in a year and a half. Libraries must hone their vision to see as far into the future as possible when choosing technology that will survive the transitions to newer and better versions. They also must choose technology that does not shut out users with older versions of systems.

The Library is aware of this tension and its sensitive to it. Obviously, within the Library of Congress, the user technology environment is controlled. But once the services move away from this controlled environment, not only is control lost but knowledge of the environment becomes sparse. Such tension may put real constraints on what services the Library can deliver via the Internet.

One resource the Library could exploit is the multimedia expertise of other organizations. Groups with digital organizational skills concerning particular collections could serve as proving grounds and educational resources for the Library of Congress. For example, the Defense Mapping Agency, the U.S. Geological Survey, and the National Geographic Society all have better map organization and digital experience than the Library of Congress. Cooperating

with others to use specialized skills and knowledge bases could save the library time and money. The Library could serve to focus the experts on a universal system for navigation of such collections.

Language Differences

The role of language has a strong effect on the successful navigation of a large database. If someone performing a search does not use the same terminology as the person indexing the material, the search might be meaningless. Medical terminology is a useful example. If a query was based on "heart disease", for instance, the system might not find the reference because it was input using a more technical medical term.

The role of language will become very important because different groups will look at a library collection from very different perspectives. People will want to interact with the computer system in their own natural languages. Even if the system moves away from key-word searching and allows full natural-language navigation, the language of a particular user group must work with the retrieval mechanisms mapped into the system. Increasing access to collections will be dependent on successful retrieval using familiar language patterns.

Session IV: Custodianship of Digital Collections

MARYBETH PETERS

Register of Copyrights, Moderator

Many people have heralded the electronic age as a way to create a paperless society and cut down on the document overload. But electronic documents have spawned the ability to save multiple versions of a work.

- Since the Library of Congress is obviously a repository, what should it be collecting?
- What should it be preserving?
- An electronic environment brings out special concerns about which version of a work should be collected. An author can change a document in a moment, and then other people down the line can also change it. Where will the Library of Congress fit into this electronic environment?

Now that publishing is moving toward digital dissemination, many works will be created electronically rather than on paper. As artists create works that are exclusively digital, what does this do to the role of the institution as a custodian of information? What happens when there is no difference between a copy of a work and the original? The work "original" will have new implications as works are created in, rather than converted to, digital form.

Copyright

Copyright law must be examined to ensure it is broad enough to cover electronic copying practices. Two issues must be considered: how does the Library manage copyright using computer systems and networks, and how does it enforce copyright law in a digital world?

The subject of scanning paper documents brings up further questions about the Library's future role. One participant suggested that libraries still will want to keep physical artifacts in addition to digital objects. For legal purposes, they will must be able to point to the original. As a record of culture, the original is essential. One participant suggested that the Library might become a registry for scanning, much as it serves as the registry for copyrights. People then would have

a central source for information on which items already had been scanned into digital form.

Royalty Payments

The issue of royalty payments must be addressed in a new way, so that authors and publishers can continue to be compensated for their work. One participant suggested that publishers are limited the number of documents they allow online because they do not want to part with their valuable products. Publishers are accustomed to getting compensated for the distribution of a work, and they are fearful that the electronic environment will make it easier for people to copy a work without payment. Many companies do not know how to deal with the rights to electronic objects, and that is impending their business. As electronic business becomes more commonplace, people will need a directory for finding the proper recipient of royalty payments. The Library is working on an electronic directory of copyright owners, but the deposit using the new electronic system is currently voluntary.

Summary Session: Steps Needed to Create a National Digital Library

JAMES H. BILLINGTON

The Librarian of Congress, Discussion Leader

The vision of a National Digital Library overlaps with many of the ideas for creating the National Information Infrastructure also known as the information superhighway. The digitization of library collections will make up one section of that information superhighway. As the Library of Congress prepares to create and contribute its portion of a National Digital Library it must articulate specific steps that will bring that goal closer.

Expert Advice

As the Library sets priorities for which items to digitize first, it must examine the goals of digitization and the audience. It also must make its potential partners aware of the subject matter that must be converted to electronic form. One conference participant suggested that the Library draw up a list of all the special collections with summary abstract descriptions and samples of the contents.

Dr. Billington said he had received a suggestion that as the Library expands its American Memory project, the data created should reflect curators' selections rather than merely data transfer. "The real value that we have to add is the curator's intimate knowledge of having worked with these collections for 30 years", he said. The online data then would be more like an anthology than a universal collection.

The Library also could make available online descriptions of the contents of collections. The availability of such synopses electronically could create further interest in the collections. Researchers would delve more deeply into the collections and request digital copies of specific works. These online directories could, in effect, serve as a menu for digitizing-on-demand.

Preservation

As the Library considers what to digitize first, one option suggested is to start with collections that are in the most urgent need of preservation because they are disintegrating. Dr. Billington pointed out that a great many materials are deteriorating at the Library (except the Gutenberg Bible, drafts of the Bill of

Rights, and selected other items, including those of vellum). If Library staff were to undertake a survey of the materials to see which collections were in the most dire need of preservation, it would take a great deal of time and could further delay the start of digitization effort.

The Library of Congress does have collections that it knows to be in critical danger of self-destructing, and digitizing those first would be helpful. But the Library has an overall plan that includes proper storage and housing, special handling of materials, and environmental controls that increase the life of objects not included in past preservation efforts.

Potential for Collaboration

The Library will need support as it makes the transition to electronic collections. Collaboration with industry could be the key to success. One participant suggested that the Library participate in video-on-demand experiments and develop a policy for making American Memory materials available to publishers who could turn them into educational media. Such a collaborate effort would increase the public's awareness of the Library of Congress as a source of knowledge. In turn, that cooperation could increase the demand for and funding of digital conversion.

Besides seeking collaboration with publishers, the Library of Congress can look to other federal government agencies as a resource for joint efforts. Digital library research projects are under way at the National Science Foundation, the Advanced Research Project Agency, and NASA. Other potential partners for collaboration include leading research and industrial organizations, state and local governments, and public libraries and school systems.

Specific collaborative opportunities that participants suggested include:

- Working with telephone companies on speech compression and recognition projects that could be applied to audiotape collections.
- Joining with a computer company in a project to make Library collections available electronically to an entire school district.
- Asking a telephone company to provide a toll-free number so elementary school children could call up and request digital files. Other corporate sponsors or volunteers could help digitize the materials and send them back to the elementary students.
- Working with teachers, textbook developers, and coursework designers to stimulate use of digital library materials in K-12 classrooms.

One participant suggested that more money should come from congressional appropriations. Just as the budget allowed for the creation of the Library's physical buildings, it could pay for the digital infrastructure. More funds will be needed to make the shift "from warehouses to disks, from corridors to networks."

Visual Interfaces

The creation of a National Digital Library provides the opportunity to take the collections of the Library of Congress out to people on other parts of the country. But along with the collections, the Library's physical space contributes to the learning experience. One participant suggested taking the physical place out to the public as well. The Library could do that through visual interfaces that allow a user to navigate through rooms and collections in the Library buildings. Such an interface would help the user understand how collections are organized in the Library and might help them find the people most qualified to answer their questions.

Visual information would capture the attention of Library users and make them more interested in exploring the Library's collections. "The more personal the experience is, the more rewarding a place it is to go", a participant said. Dr. Billington noted that the Library is planning a rotating exhibit of the treasures on the Library of Congress. As the Library moves toward a digital service environment, he said, he hopes it also can become a place that more of the public can relate to and "see" electronically.

Finding Aids

One of the benefits of the Internet and global digital environments is that a user can get access to large quantities of information from many different sources while sitting at one computer. The big problem that arises is how to find that information when the network is overloaded with files about so many different subjects. Many people say the solution would be having the Library of Congress act as a pointer to information in libraries across the country. Dr. Billington, however, questioned whether that was an appropriate role for the Library. Should the Library of Congress be a finding aid for "what's in everybody's attic and every company's archives, what's in every studio's vaults, etc.?" he asked.

If Library officials decide that is an appropriate responsibility for the Library of Congress to take on, what steps would be needed? Should the Library of Congress actively set up a system for gathering and indexing the information about other organizations' holdings? Or should the Library step back and let user need and the system evolve on their own? In one case, the Library would be passively storing bits, and in the other it would provide professional indexing

services. The consensus was that the boundaries between those two scenarios are fading. The Library will have to be flexible and find a middle ground between the two extremes.

Standards

If the Library of Congress is to be a leader in the building of a National Digital Library, it must help push for standards in the quality and processing of digital collections. The Library is not in a position to define the technical standards, but rather to push for a consensus on digitizing at the highest quality level. The consensus among participants was that the Library of Congress is in a good position to promote standards for digital interchange of information.

Where to begin

Dr. Billington asked participants what steps they thought the Library needed to undertake immediately to begin the creation to a National Digital Library. Participants said the Library of Congress should:

- Start building its digital archives.
- Strive to collect a breadth of examples of the history of America.
- Focus on a few important applications for digitized material, such as use for K-12 education, assimilation of information, and archiving.
- Set a high-quality standard for digitizing.
- Set a deadline for requiring digital copyright deposit of new material.
- Exploit the knowledge of Library staff in selecting components of the special collections for digitization.
- Develop three or four models of a digital library based on what is important in terms of policy and intellectual property considerations.

Dr. Billington and other library officials will consider these suggestions as the design for building the National Digital Library progresses.

Conclusions

The Library of Congress will be a leader in the design and implementation of a National Digital Library, but it cannot build such an ambitious infrastructure alone. Collaboration with other libraries, federal agencies, private partners, and educational institutions will be key to the success of the project. This conference

unearthed meaningful opportunities for collaboration and important matters for further consideration.

The areas that will require focused attention include:

- Standards for the creation and dissemination of digital library collections.
- The quality level at which Library materials are digitized.
- Funding for conversion of holdings to digital files.
- The limits of current technology and the potential for future advances.
- Strategies for collecting digital items.
- The development of new organizational techniques for digital collections.
- The creation of advanced navigation techniques for electronic collections.
- Intellectual property rights in a digital world.
- The potential for educational and research benefits from electronic collections.

As Dr. Billington and other Library officials imagine the libraries of the future, they will ponder these challenges and work to make the National Digital Library a resource for people across the country. The Library of Congress then truly will become "a library without thout walls".

Appendix**Conference Participants****From Industry:**

Daniel Atkins	Dean, School of Information and Library Studies, University of Michigan
Greg Blonder	Director Materials and Technology Integration Research, AT&T Bell Laboratories
Tony Dahbura	Motorola Cambridge Research Center
Bran Ferren	Senior Vice President, Walt Disney Imagineering
Henry Gladney	IBM Research Staff, Almaden Research Center
Sandra Gordon	Project Coordinator, Massachusetts Institute of Technology
Kris Halvorsen	Principal Scientist Laboratory Manager, Xerox Palo Alto Research Center
Michael Hawley	Assistant Professor of Media Technology, Massachusetts Institute of Technology Conference Co-Chair
Joe Hill	Electronic Data Systems Research
Robert Kahn	President, Corporation for National Research Initiatives
Butler Lampson	Digital Equipment Corp., Cambridge Research Lab
Michael Lesk	Executive Director, Bellcore
Clifford Lynch	Director, Library Automation, University of California
Dave Ricci	Director, Research Services, Hewlett-Packard Laboratories
Greg Riker	Director, Advanced Consumer Technology, Microsoft Corp.
Robert Sproull	Vice President Sun Lab, Sun Microsystems
James Young	Assistant to the Chairman, Electronic Data Systems Corp.

From the Library of Congress:

James H. Billington	The Librarian of Congress
Herbert S. Becker	Director, Information Technology Services, Office of the Librarian

Laura Campbell	Director, Library Distribution Service, Constituent Services
Hiram Davis	Deputy Librarian of Congress
Deanna Marcum	Director, Public Service and Collection Management I, Collection Services
Rubens Medina	Law Librarian
Marybeth Peters	Register of Copyrights
James Pruett	Chief, Music Division
Winston Tabb	Associate Librarian for Collection Services
Sarah Thomas	Director, Cataloging Collection Services
Suzanne Thorin	Chief of Staff, Office of the Librarian
Robert Zich	Director, Electronic Programs Office, Cultural Affairs Conference Co-Chair

TESORI MINIATI

Manuscripts and Early Prints
from Brescia and Bergamo
on CD ROM

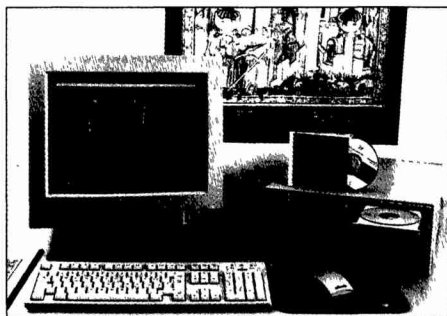
Invaluable books
from past times
are now available
on a new medium
to an even
wider public.



In cooperation with
GLASOR SNC./BERGAMO,
the
AKADEMISCHE
DRUCK- U.
VERLAGSANSTALT
presents
a series of manuscripts
and early prints
from the
Biblioteca Angelo Mai
in Bergamo
and from the
Biblioteca Queriniana
in Brescia
on CD ROM.

**Hardware and
software
requirements:**

IBM computer or
compatible PC, at
least 486 processor
and 8 MB RAM.
Windows platform,
CD ROM player,
Super VGA monitor.
Or: System Apple.



New issues
featuring further
libraries
in Upper Italy
(for instance
Biblioteca Ambrosiana
in Milan) are
being prepared.

Currently we offer 128 manuscripts and early prints
on 93 CD ROMs.

Introductory offer:
ATS 350,-/DEM 50,-/SFR 52,-



Akademische
Druck- u. Verlagsanstalt
Graz/Austria