

Digital Library

DAVID SEAMAN / JERRY GEORGE

The Digital Library Federation In America: A Portrait



David Seaman



Jerry George

Das Angebot digitaler Informationen entwickelt sich immer mehr zu einer zukunftsorientierten Dienstleistung wissenschaftlicher Bibliotheken. Dabei erweisen sich Kooperationen auf nationaler und internationaler Ebene als wichtige Bausteine sowohl für die Entwicklung allgemein anerkannter Standards und Richtlinien als auch für den effizienten Einsatz von Ressourcen.

Der nachfolgende Aufsatz widmet sich der Digital Library Federation (DLF), einem Zusammenschluss von 30 Bibliotheken und Institutionen in den USA. Als Autor dieses Porträts konnte der Direktor der DLF, David Seaman, gewonnen werden. Wir nutzen damit die Gelegenheit, einen der wichtigsten Akteure auf dem Feld der Digitalisierung selbst zu Wort kommen zu lassen und Informationen aus erster Hand zu bieten – auch wenn nicht alle Fragen deutscher Leser behandelt werden und die Innenperspektive naturgemäß deutlich durchscheint. Die Redaktion

The Digital Library Federation is a consortium of research libraries and related organizations in the United States that are pioneering in the use of electronic-information technologies. Each of the DLF's members, mostly research libraries, contributes to a capital fund and an annual budget to support the organization, which is administered by a directorate within the U.S. Council on Library and Information Resources. A Steering Committee, on which each member institution is represented, guides the work, which has focused on coordinating research, identifying standards and »best practices« for digital library development, and helping to start projects that libraries need but cannot easily develop individually. The DLF convenes task forces, issues publications, sponsors a biannual Digital Library Forum, and maintains a Web site.¹

INTRODUCTION

Visitors to the German-American Center for Visiting Scholars in Washington, D.C., the capital of the United States, may have strolled north to Massachusetts Avenue, along which many nations maintain their embassies. Visitors are unlikely to notice, on that very same avenue, a narrow building that houses, among other organizations, the Digital Library Federation (DLF). But if such visitors happen to undertake research using the World Wide Web to find digital catalogs and collections provided by major U.S. libraries, they will be profiting from the DLF's existence. For more than seven years, it has promoted pioneering work to make such resources electronically accessible. This article explains what the DLF is, how and why it began, how it operates, what activities it has undertaken, and what it plans for the future.

THE DLF: WHAT IS IT?

On 1 May 1995, directors of fourteen of the largest research libraries in the United States and two other organizations penned their names to an agreement both unusual and ambitious. It was unusual because

American libraries, which traditionally operate individually without centralized authority, agreed to a new major collaboration. It was ambitious because of what they jointly proposed to do. »We the undersigned,« they wrote,

»... agree to collaborate towards the establishment of a Digital Library Federation. The Federation's mission is to bring together – from across the nation and beyond – digitized materials that will be made accessible to students, scholars, and citizens everywhere, and that document the building and dynamics of America's heritage and cultures.«²

Impetus came from the Library of Congress, well known to scholars internationally for its extensive collections in many fields of study; it is the U.S. institution that comes closest to being a comprehensive, national library. Though created as a resource for the U.S. government's legislative branch, the Library of Congress also serves the general public. In the mid-1990s, it began developing a digital collection called »American Memory,« which now offers Internet access to more than seven million digital items.³ However, digital collections also were being created at several other American repositories, which sought to join with the Library of Congress in a »broadly collaborative undertaking.« Those besides the librarian of Congress who signed the resulting inaugural agreement to form the DLF represented the National Archives and Records Administration, the New York Public Library, twelve major American university libraries,⁴ and a private, non-profit organization in the United States called the Commission on Preservation and Access, which agreed to house the Federation and coordinate its work.

The 1995 agreement set out seven goals, which, summarized, called for

- (1) digitizing collections for a »distributed, open digital library« that would be »accessible across the global Internet«
- (2) establishing a »collective management structure«
- (3) developing a »coordinated funding strategy«
- (4) setting »selection guidelines«
- (5) adopting »common standards and best practices«
- (6) addressing issues that »may inhibit full citizen access«

(7) evaluating use of the digital library.⁵

Several months later, in 1996, the DLF formally organized.⁶ By then, the Commission on Preservation and Access had merged into the Council on Library and Information Resources (CLIR), an independent, nonprofit American organization. Through conferences, projects, and publications, CLIR helps librarians and others maintain and improve access to scholarly information for generations to come in the United States and around the world. Since 1997 CLIR has provided an administrative home to the DLF.

Operations fully began in October of 1997 when Donald Waters, formerly of Yale University, joined CLIR's staff as the DLF's first director. He oversaw the creation of a governance structure, a program plan, and the launching of several multi-institutional digital-library projects before leaving to become program director for scholarly communication at The Andrew W. Mellon Foundation, which has provided encouragement and financial assistance to CLIR and the DLF. Progress continued under Waters' successor, Daniel Greenstein, who had previously directed The Arts and Humanities Data Service in the United Kingdom. In July 2002, after Greenstein departed to become director of the California Digital Library, David Seaman, a digital library administrator from the University of Virginia, became, and remains, the DLF's director.

In its actual development, the DLF became something different from what its founders initially envisioned. The original emphasis on collaboratively creating a national digital library of Americana gave way to a broader program of helping institutions within the DLF develop their individual digital libraries, whatever the content. By the year 2000, a strategic plan for the DLF called it »a consortium of research libraries that are transforming themselves and their institutional roles by exploiting network and digital technologies.« Pursuing the plan's seven goals, DLF members have subsequently worked together to

- share and evaluate information about digital library tools, methods, practices, trends, and strategies
- stimulate digital library research and development
- respond quickly to digital library challenges as they arise
- exercise influence for the library community over a rapidly changing information landscape
- act as a catalyst in the development of innovative information services and organizations, and as an agent of learning for the profession
- attract investment in essential digital library research and development activities

– build a community of professionals appropriate to the development of digital libraries.

In 2002 the Digital Library Federation, after a substantial evaluation of its activities, concluded that its work should continue for at least five more years, until June 2007, when it is to be evaluated again.

One indicator of the DLF's utility has been the near-doubling of its membership to thirty institutions. On the list as this is written are the Library of Congress, the National Archives and Records Administration, the New York Public Library, the California Digital Library, CLIR, and libraries at these twenty-five universities: California (Berkeley), Carnegie-Mellon, Chicago, Columbia, Cornell, Dartmouth, Emory, Harvard, Illinois (Urbana-Champaign), Indiana, Johns Hopkins, Michigan, Minnesota, New York, North Carolina State, Pennsylvania, Pennsylvania State, Princeton, Southern California, Stanford, Tennessee, Texas (Austin), Virginia, Washington, and Yale. In addition, four organizations that are active in the electronic-information field have formally affiliated with the DLF as non-voting but participating »Allies«: the Coalition for Networked Information (CNI), the Online Computer Library Center (OCLC), the Research Libraries Group (RLG), and the Los Alamos National Laboratory Research Library.

Thus, the Digital Library Federation has evolved into an extensive consortium of American libraries and related agencies that are pioneering in the use of electronic-information technologies to extend their collections and services. Today, the DLF concentrates on helping its members and others by

- identifying standards and »best practices« for digital collections and network access
- coordinating leading-edge research and development
- helping start projects and services that libraries need but cannot develop individually.

HOW DOES THE DLF OPERATE?

The DLF operates through a small staff headed by a professional director, who reports to a Steering Committee on which each member institution is represented. The Council on Library and Information Resources houses the staff, provides administrative support, and collaborates in publishing reports. Major DLF initiatives, determined by the director and the Steering Committee (and a subset of it, the 5-person Executive Committee), change with needs; as some projects come to fruition or find new support, the DLF invests in others, staying flexible as a catalyst for experiment and change. Financing comes from members and grants. Member institutions make annual payments to the DLF's operating fund in amounts specified by

the DLF formally organizes in 1996

»pioneers« in the use of electronic information

initiatives change with needs

the Steering Committee. In addition, members pledge specified amounts to be paid over a five-year period to the DLF's capital fund. These funds enable the DLF to provide the following services:

- *leadership and support* for research, standards development, project start-ups, and other special initiatives
- *a twice-a-year Forum* at which staff from member institutions report developments and share experiences in developing and managing electronic resources
- *a Web site* (www.diglib.org/) for providing public access to information about digital library activities, resources, and developments, and about the DLF itself
- *three online databases* providing access through the DLF Web site to digital collections of publicly-accessible material, digital-library events, and digital-library documentation (including policies, strategies, working papers, standards, and technical documentation)
- *a Web-based DFL Newsletter* that reports on the progress of initiatives and on members' digital-library services, collections, projects, and concerns
- *publications*, printed and electronic, that report the results of research and conferences.

DLF Web site

The DLF's publications and Web site resources are available not just to members but to anyone interested worldwide.

WHAT HAS THE DLF DONE?

Some of the DLF's most important contributions have been in the development of digital-library standards and guidelines, of which the following are leading examples.

Standards and Guidelines

The Metadata Encoding and Transmission Standard (METS)

To maintain a library of digital objects, one must maintain information about those objects called *metadata*. The metadata needed for managing and using digital objects are often more extensive than the metadata used for managing collections of printed works. A library may record descriptive information about a book, but the book will not break into unconnected pages if the library fails to include a description of the book's structure, nor will scholars be unable to evaluate the book's worth if the library fails to record the kind of press on which it was printed. Metadata is essential, however, for using a book (and other library materials) in digital form. Without structural metadata, the page image or text files will be of little use, and without technical metadata about the digitiza-

tion process, scholars will not be sure how accurately the digital version reproduces the original. Also, for internal management, a library must have appropriate technical metadata to refresh and migrate the content periodically so that the resource will remain usable.

Limitations on formats for encoding descriptive, administrative, and structural metadata led the DLF to sponsor the development of METS, which stands for Metadata Encoding and Transmission Standard. The goal of the METS project was to provide an XML document format for encoding metadata needed for managing an expanded range of digital library materials, exchanging them between repositories, and providing them to users. DLF working groups have developed and refined METS XML schema and documentation.⁷

A Benchmark for Digital Master Fidelity

As digital library leaders started building online collections, they began to realize how much their individual collections could be enhanced if made interoperable with others. Online patrons could then make use of books and serial publications digitized and managed at many different sites. But to make collections interoperable, consistent formatting would be desirable in accordance with some minimum set of standards or benchmarks for digitizing items with fidelity. Out of meetings sponsored by the DLF came a set of recommended technical specifications constituting »a minimum benchmark for a faithful digital reproduction of a printed book or serial publication«. A DLF group identified several needs for additional research and established a process for wider review of the benchmark.⁸

The LIBLICENSE Model for Electronic Access

The LIBLICENSE Project arose out of recognition that new business arrangements would be needed if libraries were to obtain the right to provide scholars, teachers, students, and other patrons with access to rapidly growing numbers of electronic publications and other digital information resources not in the public domain. For print materials, libraries simply purchase copies of books and journals, which they circulate freely to their patrons. But by the mid- to late-1990s, librarians found themselves having to negotiate licenses with publishers for access rights to electronic information under copyright.

To help librarians, vendors, publishers, and others to understand electronic licensing and negotiate effective agreements, a staff team at Yale University, in consultation with librarians, publishers, lawyers, and university licensing officials, developed a model library

new business
arrangements for
electronic access

electronic-information licensing agreement, using financial support from the DLF. Additionally, the project team created LIBLICENSE software that university librarians, academic publishers, and others could freely download and use to create and customize the licenses to their own electronic resources. The project also came to include an e-mail discussion list, *liblicense-l*, with some 2,500 subscribers internationally. The model license, the software, and much helpful information about contracting for digital resources may be found on the LIBLICENSE Web site.⁹

Other Standards and Guidance

The DLF has helped produce standards and guidance in other areas as well. The DLF joined the Getty Grant Program in sponsoring ongoing work by the Visual Resources Association to compile a manual with guidelines for describing, documenting, and cataloging cultural objects and their visual surrogates. The DLF and the National Information Standards Organization in the United States are collaborating to facilitate the management of libraries' growing information resources by supporting development of a standard set of data definitions and common XML schema along with a standard format for exchanging serials subscription information. And the DLF has endorsed principles defining quality for digitized resources, published in a document entitled, »A Framework of Guidance for Building Good Digital Collections,« issued by the Digital Library Forum, a group of librarians and other experts convened by the Institute of Museum and Library Services, an agency of the U.S. federal government.¹⁰

Special Initiatives

Additionally, the DLF has undertaken special initiatives to develop tools and techniques for improving digital library collections and services. Some major examples follow.

The Open Archives Initiative (OAI)

The limitations of commercial Internet search engines, and the uneven quality of what they turn up, often make it difficult for researchers to find all material that is pertinent to their projects. The Open Archives Initiative (OAI) is an effort to decrease that difficulty.

The OAI began in a meeting in Santa Fe, New Mexico, in 1999, sponsored by the DLF and four other organizations serving libraries and publishers. The meeting produced the »Santa Fe Convention,« a technical framework for metadata »harvesting,« which the DLF subsequently explored for its potential to enable researchers to locate material on their subjects, regard-

An der Bayerischen Staatsbibliothek ist voraussichtlich im Februar 2004 die Stelle

**der Leiterin / des Leiters
der Abteilung für Handschriften und Alte Drucke
(BesGr A 16)**

wieder zu besetzen.

Die Abteilung umfasst die Sachgebiete abendländische und orientalische/asiatische Handschriften, Nachlässe und Autographen sowie Alte Drucke (bis 1700) und Seltene Drucke (einschl. moderne Libri Selecti). Auf den einzelnen Arbeitsgebieten kommt der Bibliothek Weltgeltung zu. Dementsprechend ist die Abteilung durch eine hoch qualifizierte, große Nachfrage geprägt. Für die besonderen Belange wird ein eigener Lesesaal betrieben. Eine wesentliche Rolle spielen ferner das DFG-Handschriftenzentrum und andere Erschließungsprojekte einschließlich der nationalbibliographischen Unternehmungen ISTC, VD16 und VD17. Zuständigkeiten bestehen auch für die musealen Bestände der nachgeordneten neun Staatlichen Bibliotheken in Bayern.

Aufgaben

- Leitung der Abteilung (rund 50 Mitarbeiterinnen und Mitarbeiter)
- Konzeption und Organisation der Erwerbung, Erschließung, Verwaltung und Benutzung des musealen Bestandes
- Entscheidungskompetenz in Erwerbungsfragen (Jahresetat rund 1 Mill. €)
- Organisation der einschlägigen Geschäftsabläufe
- Federführung für das Publikationswesen der Bayerischen Staatsbibliothek und der Staatlichen Bibliotheken
- Koordination und ggf. Organisation von Ausstellungen der Bayerischen Staatsbibliothek und der Staatlichen Bibliotheken
- fachbehördliche Kompetenz für Bayern in Fragen, die Handschriften und Alte sowie Seltene Drucke betreffen
- Vertretung der Bayerischen Staatsbibliothek in regionalen, nationalen und internationalen Gremien bei die Abteilung betreffenden Fragen.

Anforderungen

- abgeschlossenes historisch-philologisches Hochschulstudium (Promotion erwünscht)
- Befähigung für die Laufbahn des höheren Bibliotheksdienstes
- mehrjährige Leitungserfahrung mit flexiblen Organisations- und Betriebsstrukturen
- mehrjährige Tätigkeit im Bereich Handschriften/Alte Drucke
- Aufgeschlossenheit für digitale Dienste (möglichst erfahrungsbasiert)
- ausgeprägte IT-Kompetenz
- kooperativer Führungsstil und ausgeprägte Teamfähigkeit
- hohe soziale Kompetenz

Frauen werden ausdrücklich aufgefordert, sich zu bewerben. Schwerbehinderte werden bei gleicher Qualifikation bevorzugt.

Ihre Bewerbungen mit den üblichen Unterlagen senden Sie bitte bis zum **31. Juli 2003** an den **Generaldirektor der Bayerischen Staatsbibliothek, 80328 München**.

less of location and format, through specialized Internet portal services. Additional meetings in 2000 led to a vision statement about the convention's possibilities, a consensus in support of more work on the convention, and the formation of an OAI Steering Committee to oversee its development. With financing from the DLF and the Coalition for Networked Information (CNI), an OAI technical working group revised the Santa Fe Convention into what is now known as the OAI Harvesting Protocol.

The DLF is now supporting the testing of that protocol. The OAI Steering Committee called for a period of such experimental implementation before any further revision is made. With support from The Andrew W. Mellon Foundation, the DLF is working with the OAI towards development of a small number of Internet gateways through which users will be able to access holdings from multiple libraries as if they were part of a single collection. For testing purposes, DLF members agreed to contribute metadata from nearly 60 online digital collections representing several million unique information objects. Work is underway to make metadata available for »harvesting« by a number of specialized retrieval services.¹¹

Digital Reference Linking

Digital technology makes it possible for electronic journals to provide a service that print journals cannot – e-journals can give readers not just citations to source materials but also direct access to the cited sources themselves.

Development of the Digital Object Identifier (DOI) and of systems such as CrossRef, the first major implementation of cross-publisher reference linking, enabled scholars to navigate seamlessly from citations (whether in abstract and indexing services or in footnotes in electronic journals) to the works cited, when available in electronic form.

But initially the DOI was capable of finding only a single copy of a cited work – generally the one stored at the publisher's site. That meant that citations could not be linked to articles loaded into local systems (such as OhioLink or the University of Toronto) or into aggregator databases (such as OCLC, ProQuest, or Ebsco). Users with access to such alternate copies would be sent instead to the publisher's site and either turned away or asked to pay. The DLF worried that this limitation would become more severe as more e-print aggregations and archiving services developed – possibly even creating a monopoly for publishers, not just in the original marketing of articles, but also in their long-term accessibility.

In response, the DLF joined with others to work on

a process for locating multiple copies of articles. Out of a workshop involving representatives from CrossRef, several major scholarly and commercial publishers, the International DOI Foundation, the digital library research community, NISO, and research libraries came general agreement to permit what is now called »localization« in linking, and a commitment by CrossRef and others to develop prototypes for evaluation.¹²

The Academic Image Cooperative and ArtSTOR

The DLF has fostered creation of an electronic database of images of works of art for use by teachers and students in courses on the history of art. Planning for this scalable database began with the forming of the Academic Image Cooperative (AIC) in January 1999. With funding from The Andrew W. Mellon Foundation and support from the College Art Association, the AIC produced a prototype database and image collection in August 2000. Also, the AIC developed technical, organizational, and policy frameworks for a more ambitious, online image service – a service capable of identifying, developing, and disseminating a far larger number of artwork image collections for curricular and scholarly use.

Once planning was complete, the DLF and the Mellon Foundation began exploring how the AIC's image collection – and the know-how gained from its development – might be used in a broader undertaking. Drawing upon the DLF's work in defining viable collection policies and strategies, the Mellon Foundation moved forward in April 2001 with plans to form ArtSTOR, described as »an independent not-for-profit organization that will electronically develop, ›store,‹ and distribute digital images and related scholarly materials for the study of art, architecture, and other fields in the humanities.«¹³

Electronic Journal Archives Program

Increasingly scholarly journals are published electronically. What will it take to keep them accessible electronically in perpetuity? Can the property rights of publishers, the access responsibilities of libraries, and the reliability assurances that scholars need be reconciled in agreements to create archives of electronic journals? Seven major libraries received grants from The Andrew W. Mellon Foundation to explore solutions in a program that the Digital Library Federation helped develop and continues to assist.¹⁴

Published Reports

The DLF has extended its resources beyond its own members by publishing special reports. The following examples may be accessed electronically at no cost,

or printed copies may be ordered at nominal cost, through the Web site of CLIR (the co-publisher):¹⁵

The Digital Library: A Biography, by former DLF Director Daniel Greenstein and Suzanne Thorin of Indiana University. Drawing on a survey of DLF-member institutions, the study describes how differences in leadership, organization, and relationships with academic departments and information technology staffs have influenced the varying characters of digital libraries, identifies stages in their development, and describes probable characteristics of the future digital library.

Usage and Usability Assessment: Library Practices and Concerns, by Denise Troll Covey of Carnegie Mellon University, reports on a survey of methods developed by DLF-member digital libraries to assess the use, and usability, of their online scholarly information resources and services. It explains why and how libraries use particular assessment methods, and the strengths and weaknesses of different methods.

Scholarly Work in the Humanities and the Evolving Information Environment, by William S. Brockman of Pennsylvania State University and three co-authors, reports on a study of how humanities scholars do research, how they use technology in their work, and what kind of information environment would best support their activities in the future.

Strategies for Building Digitized Collections, by Abby Smith of CLIR's staff, synthesizes the experience of libraries in digitizing collections to make available online. Using case studies, it demonstrates that digitization programs are best developed within a library's overall collection development strategy, and inquires into whether scholarship is more usefully served by digitizing special or general collections. The report provides strategic guidance for individual institutions and general directions for the library community.

Selection and Presentation of Commercially Available Electronic Resources: Issues and Practices, by Timothy Jewell of the University of Washington, reports on how research libraries select, license, present, and support the use of commercial online materials. It deals with the cost of commercial content, assesses licensing options available to libraries, describes strategies that libraries are using to contain costs, and identifies means for integrating commercial online materials into sustainable library collections.

Building Sustainable Collections of Free Third-Party Web Resources, by Louis Pitschmann of the University of Alabama, deals with the widespread practice of developing Internet resource guides on library Web pages. The author recognizes the value such »gateways« can have while raising questions about their efficacy outside of mainstream collection development.

forming of the Academic Image Cooperative (AIC) in 1999

archiving of electronic journals

Drawing on experience internationally, Pitschmann provides guidance on using »free« Internet resources to build collections.

CONCLUSION

In 2003, the DLF begins work on a new strategic plan for the next 3–5 years, in which DLF members plan to pool content centrally in a distributed, open digital library, accessible as a unified whole to users, with files available for innovative re-use. As we plot our course for the future we draw strength from the firm basis of collaborative work we have built up since 1995. Increasingly we find ourselves able to draw together coalitions of librarians, publishers, and vendors, to tackle large-scale issues such as the future shape of the scholarly journal, and to do it across the various organizations and industries involved. We draw much of our strength from the fact that we are a small, fast-moving organization, able to respond rapidly to the dramatic changes we all see in librarianship and publishing. Finally, the DLF works because large numbers of librarians, from junior staff to senior administrators, are prepared to share their efforts, excitement and expertise for the betterment not only of their individual institution but of the digital library world in general.

¹ www.diglib.org

² Charter of the Digital Library Federation, »America's Heritage: Mission and Goals for a Digital Library Federation,« available at www.diglib.org/about/dlfcharter.htm.

³ »American Memory« resources may be accessed at memory.loc.gov/.

⁴ Columbia, Cornell, Emory, Harvard, Pennsylvania State, Princeton, Stanford, Yale, and the universities of California (Berkeley), Michigan, Southern California, and Tennessee.

⁵ For the full text of the DLF's original goals, see the 1995 Charter at www.diglib.org/about/dlfcharter.htm.

⁶ At the outset of operations, the organization called itself the National Digital Library Federation but by September 1997 no longer included the word National. On this and other points of early history, see Council on Library and Information Resources, Annual Report 1996–1997 (Washington, DC: CLIR, 1997) p. 9, available at www.clir.org/pubs/annual/annrpt96/index.html.

⁷ Information about METS is available on a Web site maintained by the Library of Congress at www.loc.gov/standards/mets.

⁸ They may be read at www.diglib.org/standards/draftbmark.htm. More information about the background of the digital master benchmark is at www.diglib.org/collections/reg/reg.htm.

⁹ www.library.yale.edu/~llicense/index.shtml

¹⁰ Detail on all these activities is available through the DLF Web site at www.diglib.org/standards.htm.

¹¹ More information is available through www.clir.org/diglib/architectures.htm.

¹² Progress is reported at the DLF Web site through www.diglib.org/architectures.htm.

¹³ More information about the Academic Image Cooperative is available from the DLF Web site at www.diglib.org/collections/aic.htm.

¹⁴ Outcomes are available at the DLF Web site at www.diglib.org/preserve/ejp.htm.

¹⁵ www.clir.org/pubs/reports/reports.htm

DIE VERFASSER

David Seaman ist Direktor der Digital Library Federation. Von 1992–2002 war er Direktor des Electronic Text Center der University of Virginia Library. Digital Library Federation (DLF), Council on Library and Information Resources (CLIR), 1755 Massachusetts Ave., N.W., Washington, D.C. 20036-2124 dseaman@clir.org

Jerry George ist Special Projects Associate der Digital Library Federation. jgeorge@clir.org