

IEEE Multimedia Systems'97

Course 2

TUTORIAL TITLE: Building and Applying Digital Libraries

INSTRUCTOR: Edward A. Fox

SCHEDULE: Tuesday, June 3, 1997, 9 a.m. - 5:30 p.m.

ACKNOWLEDGEMENTS

Sponsors and Partners: Elsevier, IBM, NSF, OCLC, SURA, U.S. Dept. of Education

Co-PIs, Students, Staff

Projects

- NSF IRI-9116991, CDA 9312611, EID-9109853, CDA-9308259
- NCSTRL (ARPA, NSF), NDLTD (SURA, FIPSE)

PART 1. Introduction

PARTICIPANTS: Level: beginner to intermediate

Participants are encouraged from those working in fields like computing, library science, information science, and publishing. Those with little background on digital libraries are invited.

DESCRIPTION:

This is part of a full-day course on digital libraries, at the end of which attendees should become able to participate in design, development, evaluation, and standardization efforts related to the global movement toward digital libraries. This session, Part 1, will focus on concepts and technology from the multimedia, information retrieval, hypertext and electronic publishing fields that relate to digital libraries (DLs) - using real case studies and examples to provide a suitable context. Projects included relate to CS (ACM literature, technical reports, courseware), material science (TULIP), and graduate education (electronic theses and dissertations), as well as the NSF/ARPA/NASA Digital Library Initiative.

Our “perspective” approach will deal with DLs regarding: user and social needs; interfaces and user interaction; architectures, components, protocols; content, publishing, and capture; and systems, engines, and operations. Issues of scalability and sustainability will be explored.

This course also can serve as a stand alone tutorial on the underlying technology for digital libraries, especially information retrieval, hypertext and electronic publishing.

PART 1. Introduction

Planned Activities: There will be coverage of the following PROJECTS:

- NETLIB (numerical analysis);
- CORE (chemistry);
- TULIP (material science & engineering, with Elsevier, OCLC);
- IBM digital libraries products and projects;
- Hyper-G/HyperWave (clients and servers);
- BEV HistoryBase;
- CS technical reports (CS-TR, WATERS, NCSTRL) and related efforts;
- CS education (ACM literature, courseware on IR, multimedia, hypertext, history);
- Digital Library Initiative (CMU, Michigan, Stanford, UC Santa Barbara, Illinois);
- ETD (electronic theses and dissertations).

PERSPECTIVES:

- user and social needs
 - Allerton workshops
 - D-Lib working group
- interfaces and user interaction
 - monitoring, analyzing, visualizing and modeling traffic
- architectures, components, protocols
 - Z39.50
 - CNRI
 - agents
 - bus: Stanford
 - systems, engines, and operations
- content, publishing, and capture
- systems, engines, and operations
- scalability and sustainability

RECOMMENDATIONS:

- Join projects: electronic theses and dissertations, CS technical reports
- Use our courseware and add to it
- Work toward “Open Digital Library”

OUTLINE:

Fig. 1: Timeline of Recent Information & DL Systems [link]

Fig. 2: NCSTRL Architecture [link]

SELECTED PAGES FROM WWW --- PROJECTS:

- * Build upon existing electronic materials
 - Netlib [link] <http://www.netlib.org/>
 - + Attribute/value search [link] http://www.netlib.org/utk/misc/netlib_query.html
- * Build upon publishers' collections
 - CORE [link] <http://ei.cs.vt.edu/~cs5604/DL/DL2.html>
 - + D-Lib Magazine, Dec. 1995, Making a Digital Library, Chemistry Online Retrieval Experiment [link] <http://www.dlib.org/dlib/december95/briefings/12core.html>
 - + OCLC [link] <http://www.oclc.org:5047/oclc/research/projects/core/>
 - TULIP [link] <http://www.elsevier.nl/inca/homepage/about/resproj/tulip.shtml>
 - + Elsevier + universities + OCLC (material science & engineering)
- * Commercial services and systems
 - IBM <http://204.146.47.71:80/is/dig-lib/> [link]
 - + case studies [link] <http://204.146.47.71:80/is/dig-lib/dlis.htm>
 - + images - QBIC [link] <http://www.qbic.almaden.ibm.com/>
 - + rights management [link] <http://204.146.47.71:80/is/dig-lib/dlfnct4.htm>
- * Enhance WWW (hypertext):
 - HyperWave [link] <http://www.iicm.edu/hyperg>
 - HyperWave server features[link]
 - HyperWave author [link] <http://www.iicm.edu/hyperwave/author>
 - HyperWave author features[link]<http://www.iicm.edu/hyperwave/author/features.html>
 - HyperWave author specs [link] <http://www.iicm.edu/.../author/specifications.html>
 - Harmony [link] <http://www.iicm.edu/harmony>
 - Harmony orientation [link]
 - Harmony screens [link] <http://ei.cs.vt.edu/~cs5604/Adv/Adv-Harmony.html>
 - Harmony information structuring [link]
 - Harmony document viewers [link]
 - Amsterdam model [link] <http://ei.cs.vt.edu/~mm/gifs/Amsterdam-hm.html>
- * Community network multimedia history
 - BEV History [link] <http://history.bev.net/bevhist/>
 - + Timeline [link] <http://history.bev.net/bevhist/historyBase/mainTimeline.html>
 - + 1992 [link], Article [link]
- * Discipline - Greek Literature [link] <http://www.perseus.tufts.edu/>
- * Discipline - Computer Science
 - Technical reports
 - + CSTR [link] <http://WWW.CNRI.Reston.VA.US/home/cstr.html>, WATERS
 - + NCSTRL <http://www.ncstrl.org/> [link]
 - * Search results [link], Search results abstract [link]
 - * Doc. thumbnails ([link]), Doc. page 1 ([link])

- * Discipline - Computer Science (continued)
 - Ptrs
 - * DLs for CS <http://fox.cs.vt.edu/DLCS.html> [link]
 - * Dienst <http://researchsmp2.cc.vt.edu:8090/> ([link])
 - Interactive Learning with a Digital Library in CS ([link]) <http://ei.cs.vt.edu>
 - * Courses ([link]) <http://ei.cs.vt.edu/courses.html>
 - * CS5604 ([link]) <http://ei.cs.vt.edu/~cs5604/>
 - calendar ([link]) <http://ei.cs.vt.edu/~cs5604/f96/Calendar.html>
 - * CS4624 ([link]) <http://ei.cs.vt.edu/~mm/>
 - + Outline ([link]) http://ei.cs.vt.edu/~mm/s96/sspace/Outline_424.html
 - + Projects ([link]) <http://ei.cs.vt.edu/~mm/s97/sspace/Projects.html>
 - + Virtual Computer History Museum ([link]) <http://video.cs.vt.edu:90/history>
 - Welcome, Register, Lobby, Timeline of Events
- * DLI [link] <http://dli.grainger.uiuc.edu/national.htm>
 - + DL Info,Resources [link] <http://interspace.grainger.uiuc.edu/~bgross/digital-libraries.html>
 - + DLI Publications [link] <http://dli.grainger.uiuc.edu/pubsnatynch.htm>
- CMU (Carnegie Mellon U.)
 - + Informedia [link] <http://www.informedia.cs.cmu.edu/research/>
 - + NetBill [link] <http://www.ini.cmu.edu/NETBILL/>
- Michigan
 - + Agents [link] <http://ai.eecs.umich.edu/people/jmvidal/papers/tpa/node2.html>
 - + Groups [link] <http://http2.sils.umich.edu/UMDL/teams.html>
- Stanford [link] <http://www-diglib.stanford.edu>
 - + Testbed Development [link] <http://walrus.stanford.edu/~testbed/>
 - + Info finding [link] <http://www-diglib.stanford.edu/diglib/pub/infofinding.html>
 - + User interfaces [link] <http://www-diglib.stanford.edu/diglib/pub/userinterface.html>
 - + DLITE (task env) [link] <http://www-diglib.stanford.edu/diglib/cousins/dlite/>
 - + DLITE comps [link] <http://www-diglib.stanford.edu/.../dlite/component-list.html>
 - + DLITE screens [link] <http://www-diglib.stanford.edu/.../dlite/screen-shots.html>
- UCSB (Santa Barbara)
 - + tutorial [link] <http://alexandria.sdc.ucsb.edu:3366/doc/tutorial/index.html>
 - + 1996 ([link]) <http://alexandria.sdc.ucsb.edu/public-documents/annual-report/>
- UIIC (Illinois)
 - + interspace [link] <http://www.grainger.uiuc.edu/dli/infrastr.htm>
 - + semantics [link] <http://www.grainger.uiuc.edu/dli/semantic.htm>
- * Genre - ETDs - electronic theses and dissertations - <http://etd.vt.edu/etd/> [link]
 - Submission form [link] <http://scholar.lib.vt.edu/cgi-bin/etd.cgi>
 - Approval form [link] <http://etd.vt.edu/etd/students/approval.html>
 - Letter to students [link] <http://etd.vt.edu/etd/students/students.html>
 - Standards [link] <http://etd.vt.edu/etd/students/formats.html>

Introduction

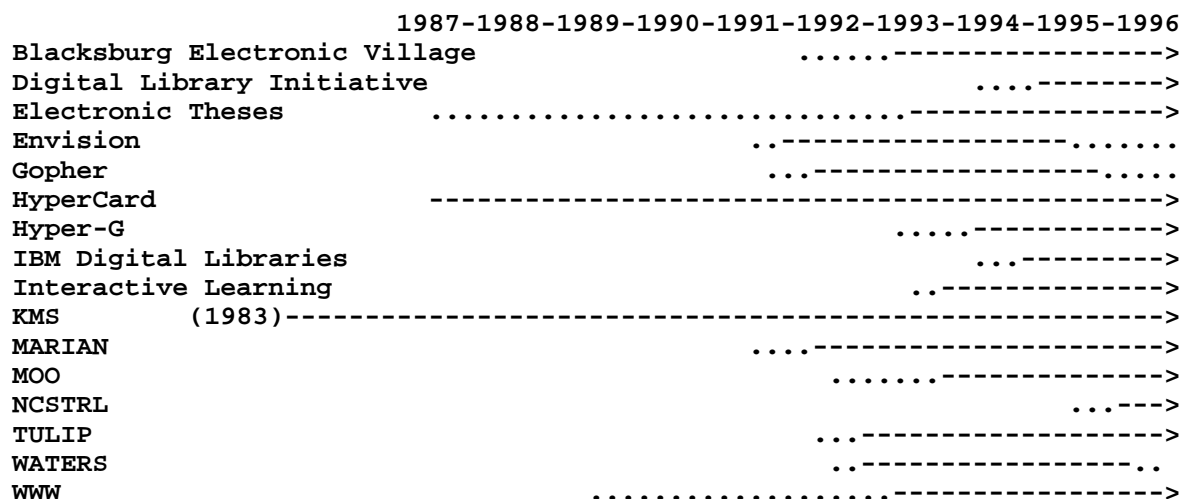


Figure 1: Timeline of Recent Information and Digital Library Systems



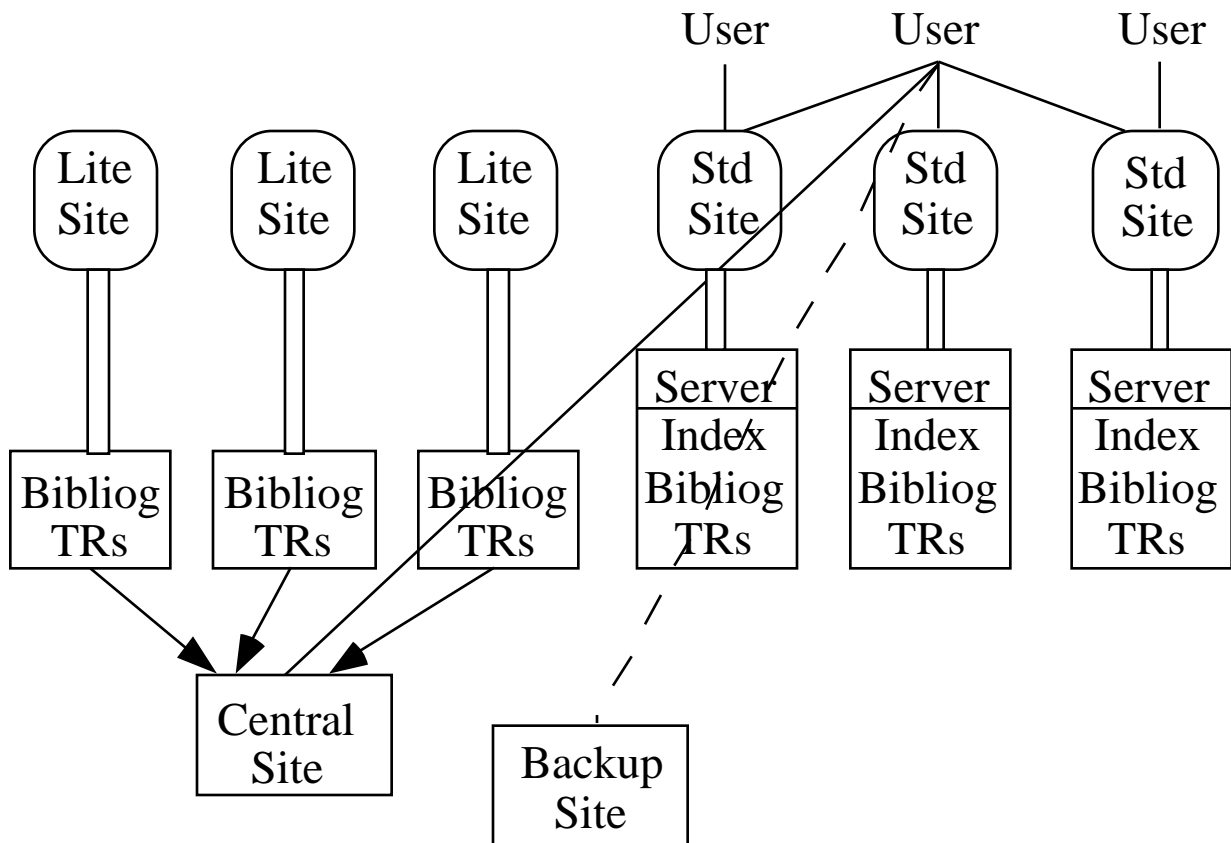


Figure 2: NCSTRL Architecture





Netlib Repository at [UTK](#) and [ORNL](#)

Netlib is a collection of mathematical software, papers, and databases.

There have been [17,998,523](#) requests to this repository as of Thu May 29 02:23:11 EDT 1997 .

Software, papers, etc.

- [Browse](#) the Netlib repository
- [Search](#) the Netlib repository

What's New?

- [Added contributions directory to the MPI directory](#)
- [BenchWeb](#)

Services provided at Netlib

- [Conferences Database](#)
- [Java Version of Linpack Benchmark](#)
- [Numerical Analysis Net \(NA-Net\)](#)
- [Performance Database Server](#)
- [Top500 Supercomputer Sites](#)

Related efforts

- [BenchWeb](#)
- [HPC-Netlib](#), high performance branch of Netlib
- [National High-Performance Software Exchange \(NHSE\)](#)
- [Parallel Tools Library \(PTLIB\)](#)
- [StatCodes](#) at Penn State, statistical source codes and packages of use to physical scientists

Information about Netlib

- [Frequently Asked Questions about Netlib \(FAQ\)](#)
 - [Netlib Editors](#)
 - [Netlib Mirror Sites](#)
 - [Netlib Server Statistics](#)
-

Search Netlib

Helpful searching tips:

- You can do a fielded search by using the syntax :
fieldname=whatever
See below for a list of field names.
- Another good way to find the software you need is to use the [GAMS](#) class hierarchy.
- You can use [freeWAIS-sf query syntax](#).

Field names:

(global indicates that the field is included in the global index that will be searched if no attribute name is specified)

- **file** -- any portion of the pathname for a regular file
- **lib** -- any portion of the pathname for a directory
- **for (global)** -- problem solved or description
- **gams** -- [GAMS](#) class
- **prec** -- Fortran precision (single, double, complex, or doublecomplex)
- **title (global)**
- **alg** -- algorithm or method
- **by (global)** -- author (name <email>)
- **keywords (global)** -- terms as would be drawn from a subject thesaurus
- **lang** -- programming language

Search Examples:

1. To search for single precision routines in the lapack directory that do Schur factorization:

```
file=lapack and file=single and Schur
```

(since the lapack single precision routines are in the lapack/single directory)

2. To search for curve fitting or gams class E1 and its subclasses:

```
(curve and fitting) or gams=e1*
```

3. To do a literal search for 'cosine transform':

```
'cosine transform'
```

Digital Libraries - Example: The CORE Project

Some digital libraries have been developed for a profession. The CORE Project is such an effort, for the field of chemistry. It involves the major US publisher and information provider of chemistry information, the American Chemical Society, and its subsidiary, Chemical Abstracts Service.

Statistics regarding CORE Digital Library:

- Pages: 430K (now 375K)
- Extracted Graphics: 387K
- Articles: 82K
- Gbytes Page Images: 50
- Gbytes Text: 4.4
- Gbytes Graphics: 6
- Gbytes Index: 11
- Scanning from: paper, microfilm
- High Resolution (to print): 300dpi (2560x3328) B&W
- Low Resolution (to display): 100dpi (856x1109) grey scale
- Conversion of Figures: extraction
- Conversion of Text: typesetter tapes to SGML
- search engine: OCLC's Newton
- Interfaces: OCLC's SCEPTER, Bellcore's Pixlook



D-Lib Magazine, December 1995

Project Briefings and Updates

Making a Digital Library

The Chemistry Online Retrieval Experiment A Summary of the CORE Project (1991-1995) December 1995

Contributed by:

Richard Entlich, Cornell University

Lorrin Garson, American Chemical Society <http://pubs.acs.org>

Michael Lesk, Bellcore

<http://community.bellcore.com/lesk/home-page.html>

Lorraine Normore, Chemical Abstracts Service

Jan Olsen, Cornell University

Stuart Weibel, OCLC <http://www.oclc.org:5046/~weibel>

The CORE project was an electronic library prototype of primary journal articles in chemistry, containing about four years of twenty primary journals published by the American Chemical Society (about 400,000 pages). CORE included both scanned images and an SGML (Standard Generalized Markup Language) marked-up version for on-the-fly rendering for screen display. Each page was scanned and segmented, with graphical units isolated and linked to figure references in the articles. The original machine-readable typography was converted to SGML format and the results were used to build databases with indexes for full-text Boolean searching.

Each page image was stored as a 300 dpi bitonal image for printing,

and 100 dpi greyscale for screen display. All text data and the most recent page images were available on Unix-based magnetic storage at any given time, with additional (older) page images stored on a WORM (Write Once, Read Many) jukebox.

Complex scientific material (superscripts, tables, equations, special fonts and symbols, etc.) presents substantial problems for representation and display, especially when the material is being converted from previously published information, as were these journals.

The tasks of building and maintaining electronic journal databases remains formidable (especially if conversion from older formats is involved). However, experiences with chemists in this project suggest that electronic publishing will be popular with scholars, even though there remain significant disadvantages and impediments to adoption.

Analysis of user studies and transaction logs is ongoing and will be submitted for publication in the near future.

Further information on the CORE Project can be found at:
<http://www.oclc.org:5047/oclc/research/projects/core>

Acknowledgments: The CORE project thanks Sony of America, Digital Equipment Corporation, Sun Microsystems, and the Cornell Theory Center (which receives major funding from the National Science Foundation, and New York State, and additional funding from ARPA, the National Institutes of Health, and IBM Corporation).



hdl://cnri.dlib/december95-briefings.2

The CORE Project: Overview

The CORE project is an electronic library prototype that provides networked access to the full text and graphics content of the American Chemical Society journals and associated Chemical Abstracts Services indexing since 1980 (some 250 journal years of data). The database is coded in SGML (Standard Generalized Markup Language) which was translated from the original typography codes, captures the structural richness of the original document and provides flexibility for indexing, searching and display. The prototype provides a full-scale laboratory environment in which to explore issues of database structure, user interface capabilities, and information retrieval questions on a large, real-world scholarly electronic journal database. The complete database, representing more than 600,000 pages of full text and graphics, will be available at Cornell University in late 1994. The major contributors of this electronic library project include:

- Cornell University (Mann Library)
- OCLC
- Bellcore
- American Chemical Society
- Chemical Abstracts Services

Relevant publications

- The CORE Project: Technical Shakedown Phase and Preliminary User Studies
- The Design and Implementation of XSCEPTER, an X-Windows Graphical User Interface to the CORE project

Some Images of XSCEPTER

XSCEPTER provides "on-the-fly" formatting of SGML as defined by configurable style guides and a DTD, to provide rapid display of scholarly data. The XSCEPTER interface is coupled with NEWTON, OCLC's proprietary search engine, to provide navigational capabilities of the CORE collection.



XSCEPTER main window



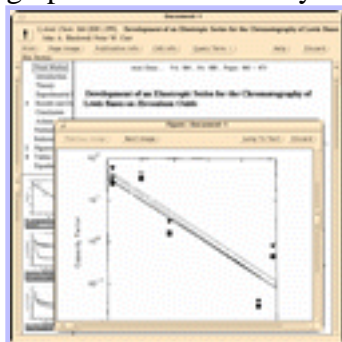
Find Box to allow indexed full boolean searching of the CORE database.



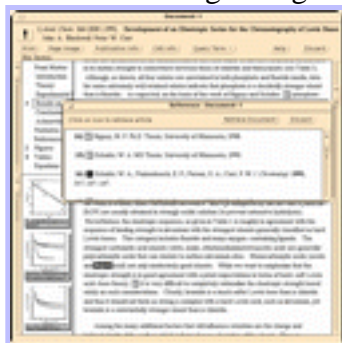
Result hit-list of database search



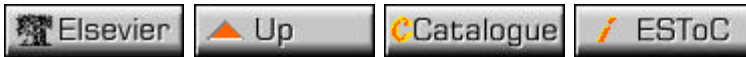
XSDOCUMENT view of article including dynamic table of content, table of graphics and "on-the-fly" rendering of SGML text.



Selection of image from table of graphics allowing hyperlink to referenced text or animation through image list.



Hyperlink of bibliographic reference to either location current document or retrieval of actual citation from database.



TULIP The University Licensing Program

When you scroll further down this page you'll find

- [Introduction](#)
 - [The TULIP Final report](#) [Now reprinted because of popular demand]
 - [TULIP Newsletters](#)
 - The [Journal Titles](#) in TULIP
 - The [Universities](#) involved in TULIP
 - The [Anonymous FTP](#) facility for TULIP
 - [Contact information](#)
-

Introduction

TULIP is a cooperative research project testing system for networked delivery and use of journals, performed by Elsevier Science and [nine Universities](#) in the USA. The participants set three objectives at the outset:

Technical

To determine the technical feasibility of networked distribution to and across institutions with varying levels of sophistication in their technical infrastructure. "Networked distribution" means sending the information both across the national Internet and over campus

networks to the desktops of students and faculty. Elsevier will deliver the journal information to participating universities in standard formats. The universities will incorporate the information in local prototype or operational systems. A wide variety of delivery alternatives, search and retrieval systems and print-on-demand options will be compared.

Organizational and economic

To understand, through the implementation of prototypes, alternative costing, pricing, subscription and market models that may be "viable" in electronic distribution scenarios; comparing such models with existing print-then- distribute models; and understanding the role of campus organizational units under such scenarios. The overall goal is to reduce the unit cost of information delivery and retrieval. "Viable" means economically and functionally acceptable to all parties.

User behaviour

To study reader usage patterns under different distribution (technical, organizational and economic) situations. Improvement in the functionality of the information, whether as to article structure or retrieval tools, will also be considered. Certain data will be collected uniformly at all sites for analysis in the aggregate and for comparison among different systems.

Click [here](#) to return to top of information

The TULIP Final report

The [final report](#) for the TULIP project is currently available, both in HTML and in a printed version.

The HTML version

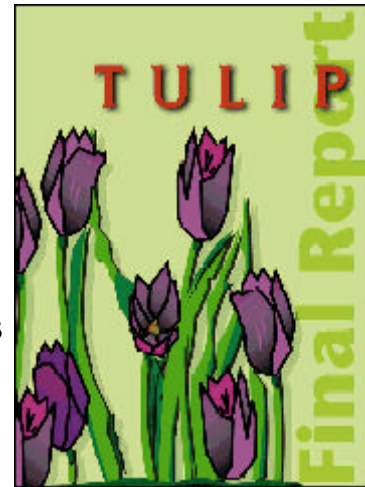
For easy printing of the entire report by your Webbrowser, it is divided in eight files

1. The [top level document](#) including the [Table of Contents](#), the [Executive Summary](#) and the [Introduction](#) (23 Kilobytes)
2. [Chapter I. Description of the project and participants](#) (40 Kilobytes)
3. [Chapter II. Technical aspects](#) (75 Kilobytes)
4. [Chapter III. Promotion](#) (9 Kilobytes)
5. [Chapter IV. User behavior](#) (45 Kilobytes + artwork 380 Kilobytes)
6. [Chapter V. Organizational and economic issues](#) (32 Kilobytes)
7. [Chapter VI. Implications of the TULIP project for the future of the development of digital libraries](#) (19 Kilobytes)
8. [The HTML document with appendices I through XIV](#) (35 Kilobytes). Please note that some appendices are not yet finished.

The printed version

We have a limited supply of the paper version of the TULIP Final Report available. Please fill in the [order form](#) to receive a free copy of the TULIP Final Report.

NOTE: Because of popular demand the stock was depleted earlier than anticipated. We have arranged a reprint



which is now available. We regret any inconvenience or delays this may have caused.

Click [here](#) to return to top of information

TULIP Newsletters

The following TULIP Newsletters are available for browsing

- [TULIP Newsletter no. 7 - July 1996](#)
- [TULIP Newsletter no. 6 - May 1995](#)
- [TULIP Newsletter no. 5 - September 1994](#)
- [TULIP Newsletter no. 4 - April 1994](#)
- [TULIP Newsletter no. 3 - January 1994](#)
- [TULIP Newsletter no. 2 - Augustus 1993](#)
- [TULIP Newsletter no. 1 - November 1992](#)

Click [here](#) to return to top of information

Journals in TULIP in the field of Material Science

The participating universities have in common strength in the physical and engineering sciences. In looking within these disciplines for a target area, we wanted a field in which the researchers were comfortable with computer applications and had a higher than average installed base of workstations. An obvious choice might have been computer science itself, but we felt these users would be so atypical in their computer facility as to make it hard to generalize results to other disciplines. Materials science provided a field in which there was both a sufficiently large corpus of frequently-cited material within one publishing company and interested faculties. Therefore [83 journal titles](#) were chosen from the collection of Elsevier Science journal titles.

Click [here](#) to return to top of information

Universities participating in TULIP

- **University of California** (all campuses)
 - Berkeley
 - Davis
 - Irvine
 - Los Angeles
 - Riverside
 - Santa Barbara
 - Santa Cruz
 - San Diego
 - San Francisco
- [Carnegie Mellon University](#) (Pittsburgh, PA)
- [Cornell University](#) (Ithaca, NY)
- [Georgia Institute of Technology](#) (Atlanta, GA)
- [University of Michigan](#) (Ann Arbor, MI)
 - A [demo version](#) of the prototype of Michigan's library system based on World Wide Web
- [Massachusetts Institute of Technology](#) (Cambridge, MA)
- [University of Tennessee](#) (Knoxville, TN)
- [Virginia Polytechnic Institute and State University](#) (Blacksburg, VA)
- [University of Washington](#) (Seattle, WA)

Click [here](#) to return to top of information

TULIP through Anonymous FTP

TULIP has a number of files available for download through Anonymous FTP. Connect to <ftp.elsevier.nl>, enter "anonymous" as

username and enter your email address as password. Change into the [/TULIP](#) directory and read the [Readme.txt](#) file.

Click [here](#) to return to top of information

For general information on the TULIP Program

Project Manager

Jaco Zijlstra (Singapore), voice +65-434 3714, fax +65-337 2230,
e-mail jaco@pacific.net.sg

Project Leader

Karen Hunter (USA), voice +1 212-633 3787, fax +1 212-633 3764,
e-mail K.Hunter@Elsevier.com

Technical Coordinator

Paul Mostert (The Netherlands), voice +31 20-485 3574, fax +31
20-485 3706, e-mail P.Mostert@Elsevier.nl

Mail Address

Elsevier Science
655 Avenue of the Americas
New York, NY 10010
United States of America

Click [here](#) to return to top of information

Last modified on July 15, 1996 by [Paul Mostert](#)

© [Copyright](#) 1997, Elsevier Science, All rights reserved.

IBM Digital Library

Welcome to IBM Digital Library. You are about to experience one of the most innovative and revolutionary of IBM's wealth of exciting solutions. Tomorrow's digital asset management system is here today and you can be a part of it.

IBM Digital Library is the digital age library for your multimedia content. It is a comprehensive product capable of creating customer specific solutions. Whether it's video, audio, images, or text, IBM Digital Library transforms multimedia assets into a digital form which can be distributed over public or private networks, like the Internet and your corporate intranets, to users around the world.

It's an open, scalable solution, and at the core of its architecture is a library server which manages the catalog information and provides secure access to objects held in the collection.

We invite you to take a look at IBM Digital Library; the product, architecture and industry solutions. You'll see why IBM Digital Library is revolutionizing the way you'll do business with your multimedia assets.

[▶ IBM HOME](#) [▶ SUPPORT](#) [▶ EMPLOYMENT](#) [▶ CONTACT IBM](#) [▶ LEGAL](#)

©1997 IBM Corporation



IBM Digital Library

Explore IBM Digital Library On A Case By Case Basis

Taking the initiative to build a solution using IBM Digital Library requires foresight, strategic acumen and trust in the technology that supports the enterprise. Arranged here by categories ([Media](#), [Higher Education](#), [Government](#) and [Cultural Institutions](#)) are case histories of several IBM Digital Library solutions. No two are exactly alike, yet all share challenges that IBM Digital Library's end-to-end capabilities are uniquely qualified to handle.

Media and Entertainment

The media industry includes a broad assortment of companies-from entertainment and broadcast to publishing and sports-all of which can benefit from investing in digital libraries. They all create, produce and distribute "content". Many take their original creative or informational assets, and re-purpose or repackage the content. For instance, entertainment companies who create, produce and distribute live action and animated films, music, live entertainment and other attractions might repackage the content through television, theme parks, interactive software and consumer products. Sports organizations who stage and distribute sporting events also redistribute their content and brands which were developed by the sports activities. In order to effectively implement this process of adding value to products, the content must be easily stored, retrieved

and distributed, regardless of its media format. This is why a digital library is the necessary foundation technology for the future success of all media and entertainment companies.

Following are some applications of IBM Digital Library in the media industry.

- Video or Audio Asset Library-Video or audio objects from clips to complete features can be managed and archived. An example would be a news archive/library, where raw and on-air news items, support research and supply clips are digitally managed, along with all the information about the video object (title, subject, date(s) aired, production information, credits, length, description, person(s) pictured in the clips, etc.) Another example would be a digital library to manage TV ads at a central site for network distribution.
- Content Creation Library/Digital Studio Backend-The library would manage objects representing individual frames in a live action or animated feature. The objects could even be the elements of a composite frame. The library could associate the information about the feature from script, through production notes and costs, to the final version of the frame element, frame or scene.
- Content Creation Library/Digital Publishing-Similar to the Digital Studio Backend, this variation would manage and archive the content from concept through final form for either print or Internet publishing, including an image of the final pre-press page.
- Interactive Software Studio Library-This library would support all the media formats (video, audio, still image, graphics and text) employed in the production of interactive software. It would manage the media objects and their various versions throughout the production process, retaining searchable information about the objects.
- An Enterprise Digital Content Library-A large media organization with television, radio, print and/or online outlets can centralize the gathering and archiving of content so that all outlets had access to the total raw feeds and the elements already produced elsewhere in the enterprise.

[CareerPath.com](#)

A new Internet employment service from CareerPath.com provides free access to more than 100,000 opportunities across the country.

[EMI's KPM Music Library](#)

Among the world's largest suppliers of music (and other media) for producers of movies, TV shows, advertising and presentations, EMI Music Publishing has opened the vaults of KPM's Music Library through IBM Digital Library and IBM's partner Multimedia Archive and Retrieval Systems plc.

[Institute for Scientific Information](#)

After searching 6 months for a technology partner to build an "electronic library system," the Institute for Scientific Information chose IBM Digital Library for good reasons. ISI's Electronic Library Pilot project contains over 1,350 journals including tables of contents, bibliographic data and abstracts.

Higher Education

The paradigm for the educational system is about to change drastically. Digital libraries will be the gateway for everyone, everywhere to gain access to the knowledge they seek regardless of where the physical sources exist. Following are some examples of applications in the area of Higher Education.

- Interactive, multimedia courseware - Institutions will be able to manage large digital collections of educational material and courseware enabling distributed learning and on-demand applications. Professors can construct course material (each with their own tailored "multimedia text book") from objects stored in the digital library. Students will be able to electronically access this material anywhere, anyplace, anytime.
- Capture, index and store content - Multimedia content from rare collections and other sources can be stored for easy access and preservation.
- Extension of traditional library catalog to digital content - Campus

libraries have always offered traditional bibliographic content. Digital libraries will enable patrons to access new types of resources, ensuring the security of information and maintaining copyright protection priorities of librarians.

[Case Western Reserve University](#)

In a joint effort with IBM, Case Western Reserve University's faculty, librarians and information service staff undertook a project to identify the critical needs and components for developing a digital library. A big part of the success story is IBM's rights management technology, which administers the terms and conditions for use of copyrighted materials.

[Indiana University School Of Music](#)

Indiana University's Variations Music Information System, created with IBM Digital Library, puts an entire music library on-line. It allows the music students to immerse themselves in their studies wherever they are, whenever they want.

[Marist College](#)

With 100,000 objects currently digitized and stored, Marist College anticipates adding an average of 10,000 more per month. Within the next four years, Marist will have several million objects in their digital library. This supports the new paradigm for learning, providing access to learning archives, at higher educational institutions.

[Virginia Commonwealth University](#)

Virginia Commonwealth University recognizes the emerging need to make education available to students anytime, anywhere. It is using IBM Digital Library to support their vision, starting with an on-line Pharmacy Case Study course. In addition, the Radiology Department of their medical college is using IBM Digital Library to develop an image transmission system to make digitized MRIs accessible to the medical community.

Government

Government organizations are looking for ways to provide better

service to their citizens. Through digital library technology, agencies can provide electronic access to documents previously unavailable to the general public, either because they are too fragile, or because it is too expensive to administer physical public access.

The following solutions can be applied to this industry.

- Virtual libraries - Agencies such as the Library of Congress, can have their collections scanned and put on the web, to be accessed by schools as well as the general public. Various collections can be linked to one another. A "user" doesn't need to know the actual physical location of any collection.
- Training - Government institutions have the need to efficiently educate their employees. IBM Digital Library technology can be used for computerized training of employees, contractors and the public to enhance the services that an agency provides.
- Information access - Digital libraries can be used to make documents and forms easily accessible to the public via the Web.

Cultural Institutions

Imagine reading the Dead Sea Scrolls. Or the Rosetta Stone. Until very recently, only the privileged few among the world's scholars could view the great artifacts of human history. With IBM Digital Library, the great works of mankind can be opened up to the world. The opportunity to preserve for future generations the knowledge of antiquity has arrived -- IBM Digital Library.

Following are some examples of IBM Digital Library solutions.

- Preservation and archiving of rare artifacts-Limited facilities and the deterioration of old and rare documents have resulted in many of the world's treasures being locked away in vaults for only a few people to study. Using IBM Digital Library technology, high quality digital images can be accessed via networks worldwide, opening up the chance for scholars and the public everywhere to study and research valuable exhibits.
- Rights management of collections-Sometimes collections are

donated to an institution with the stipulation that there will be restricted access. Digital libraries offer rights management solutions to address multiple categories of use for various collections, allowing institutions to take full advantage of entering the digital world.

[Archivo General de Indias](#)

To better serve researchers and to preserve its archives of 90 million pages of historical materials documenting the Spanish conquest of the new world, Archivo General de Indias joined IBM Spain and the Ramon Areces Foundation in creating a digital library. Currently, more than 9 million pages may be searched and accessed on 40 IBM workstations.

[Franklin D. Roosevelt Presidential Library](#)

IBM, in a joint project of Marist College, is using IBM Digital Library technology to digitize and manage the content of the FDR Presidential Library located in Hyde Park, New York.

[The Lutherhalle Wittenberg Museum](#)

With the Help of IBM Digital Library, the museum has embarked on an ambitious project--the transfer of 2.5 million objects to electronic format.

[The Vatican Library](#)

An IBM Digital Library project of monumental proportions, the Vatican Library holds over 150,000 manuscripts, including the oldest known manuscript of the Bible from 350 A.D., and 1.5 million books including 8,000 published during the first 50 years of the printing press.

[▶ IBM HOME](#) [▶ SUPPORT](#) [▶ EMPLOYMENT](#) [▶ CONTACT IBM](#) [▶ LEGAL](#)

©1997 IBM Corporation



The QBIC Project



© 1995 IBM Corporation



This site received a 4 star rating from McKinley Group's editorial team.

QBIC^(TM) -- IBM's Query By Image Content

On-line collections of images are growing larger and more common, and tools are needed to efficiently manage, organize, and navigate through them. We have developed a system called QBIC which allows queries of large image databases based on visual image content -- properties such as color percentages, color layout, and textures occurring in the images. Such queries use the visual properties of images, so you can match colors, textures and their positions without describing them in words. But, for example, a query using a red car on a green background will match a red bird on a similar green background. So we often combine content based queries with keyword and text predicates to get powerful retrieval methods for image and multimedia databases.

NEW QBIC is available for download with a free 90 day trial license. The download package includes the image indexing and search engine (for AIX, Linux, Windows NT/Windows95, and OS/2), a Web front end, APIs for imbedding QBIC in other applications or extending QBIC with new query functions, and even a sample image collection. You can download it from [IBM software download site](#).

NEW Examples of sites using QBIC include:

- [The Art and Art History QBIC project at UC Davis](#).
- [Imagebase at the Fine Arts Museums of San Francisco](#).
- [Image collections from the French Ministry of Culture](#).

For full use licenses, contact drewc@ibm.net. Also, check out QBIC's availability in the [DB2 Image Extenders](#), which are components of IBM's scalable, multimedia, Web-enabled [DB2 Universal Database](#).

To try the World Wide Web QBIC search engine on a set of approximately 1,900 images, called our "demo" catalog, press the "Run Demo" button below. Other catalogs may be available in the future. You need an HTML 3.0 capable browser.

Catalog name:

If you have any comments on this on-line demo, or would like to contact the QBIC group, write us at: qbicwww@almaden.ibm.com

If you would like to be on our mailing list, please enter your name in the following box and press Enter:

Your e-mail address:

Other links related to QBIC:

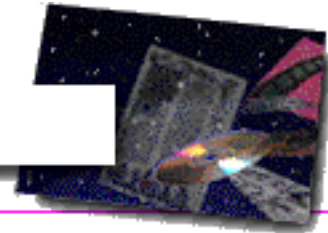
- [DB2 Extenders, which now include QBIC technology in DB2 through the Image Extender.](#)
- [Ultimedia Manager 1.1, a stand-alone product that incorporates QBIC technology.](#)
- [IBM Digital Library - Related technologies for information management.](#)
- [IBM MediaMiner - Software tools for building multimedia information retrieval and mining applications.](#)
- [Technical paper requests on QBIC.](#)

[[IBM home page](#) | [Order](#) | [Search](#) | [Contact IBM](#) | [Help](#) | [\(C\)](#) | [\(TM\)](#)]

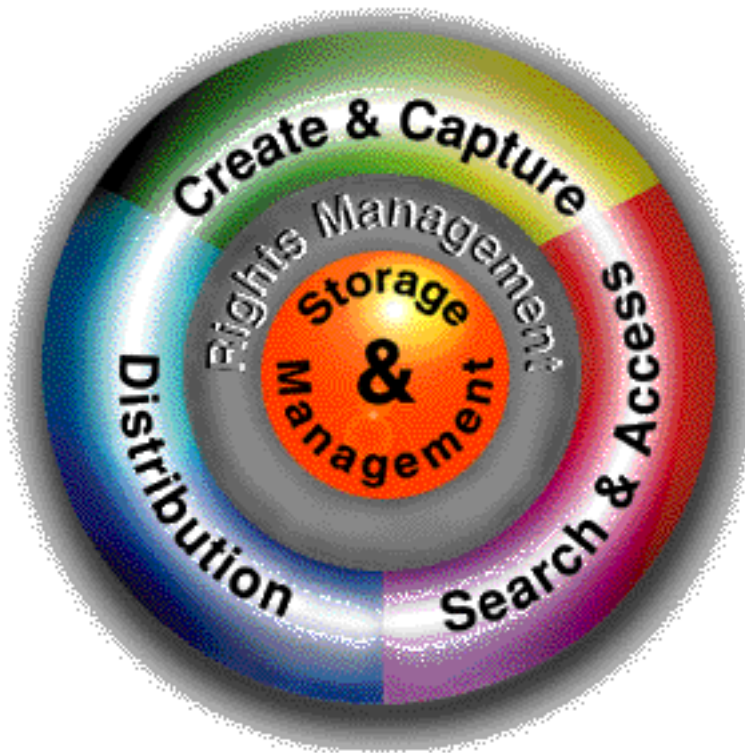


© 1995 IBM Corporation

Software



IBM Digital Library



Rights Management: New technologies bring new opportunities, but not without risk

The definition of content ownership is not universal. Copyrights may be protected in one country and practically ignored in the next. So how do you protect your intellectual property rights in a digital library that anyone with a PC can learn to use in a few minutes? Rights Management is the answer. IBM has concentrated on Rights Management as a key factor in creating IBM Digital Library and allowing you to develop a full digital library solution to meet your business needs.

The challenge is to provide ease of use, privacy, content integrity and cost utility for users while establishing bulletproof Rights Management solutions for content owners. To meet that challenge, IBM Digital Library incorporates the rich legacy of IBM's decades-long innovations

in networking security and transactions technology.

The Legacy At Work

Compare using a Digital Library to making a withdrawal from an automated banking machine. A user is identified, enters a password, requests information or an object, and the system checks the request against the user's eligibility. Upon approval the requested objects can be watermarked to deter illegal duplication. And the entire process is conducted in a secure environment without the threat of intrusion. Meanwhile, compensation for the value added to the information or object is duly processed.

Did you know that the vast majority of the world's automated banking systems are built on IBM software and technology? Trust and reliability are what make legacies, and the Rights Management functions of IBM Digital Library have inherited both.

Signed, Sealed, Delivered

IBM Digital Library can authenticate original media -- photos, manuscripts, audio, video, film and pictures -- by using electronic signatures. Digital content can be recognized as authentic with these signatures.

Watermarks, a form of electronic signature currently in use throughout several IBM Digital Library solutions, are encoded onto photos, films, videos and manuscripts. Visible watermarks can be graphically representative of a content owner's identity, like a logo or crest. Watermarks are sophisticated identifiers that inhibit the

misappropriation of content owners' assets while assuring users' confidence in the authenticity of the content.

IBM infoMarket Search service represents a giant leap forward in Network-centric computing for both content owners and users. The infoMarket service enables users to search simultaneously available network databases (private, public or both). For content owners, infoMarket provides its Plug-N-Publish® toolkits and the Cryptolope®, an encryption-protected "envelope" that can travel on public networks.

Anybody who wants to open a Cryptolope to read its contents must use a key to unlock it. Users can preview a Cryptolope's contents, then decide whether to pay for the key. For sensitive content needing increased security, a Cryptolope may require several keys. And a Cryptolope can travel on networks with only the intended recipient being aware of its presence. All the while, the infoMarket service keeps impeccable records of rights payments transacted.

IBM Digital Library follows through for all content owners. With the rise of multimedia, the work of several content authors is contained within a single media object. For instance, a digitized document might contain a photo, an illustration, a page of text and some music---each authored by a different person who should be compensated. Furthermore, a customer accessing this document might not be required to purchase the entire work, but rather just the individual section that is accessed. The Rights Management capabilities of IBM Digital Library offer discreet recognition for each content author, keeping track of who gets paid for what.

IBM Digital Library provides trusted means for protecting and managing the rights of content owners. Rights Management issues impacts every aspect of IBM Digital Library -- Create & Capture, Storage & Management, Search & Access and Distribution. Use The Wheel at the top of this page to continue exploring IBM Digital Library technology.



**Digital
Library Home**



**Request
More Info**



**Technical
White Papers**



**Software
home**



**Short
cuts**

[[IBM home page](#) | [Order](#) | [Search](#) | [Contact IBM](#) | [Help](#) | [\(C\)](#) | [\(TM\)](#)]

[HOME](#) | [SEARCH](#) | [OPTIONS](#) | [IDENTIFY](#) | [EDIT](#) | [ANNOTATE](#) | [HELP](#)

New Features of Hyperwave 2.0



What do you want to serve today...?

HYPER-G NOW HYPERWAVE

NEXT GENERATION WEB SOLUTION

[HyperWave Server](#) is advanced web server (WWW) technology, based on an **object-oriented database** which was developed especially for hypermedia document management. HyperWave guarantees **automatic hyperlink consistency** and supports hyperlinks to and from multimedia documents, **full text retrieval** and a sophisticated **access control** system with user management and user groups.

HYPERMEDIA AUTHORIZING UTILITIES



[HyperWave Author](#) - code-named "[Harmony](#)" for the UNIX version and "[Amadeus](#)" for the Windows version - accesses HyperWave servers across the Internet or any internal TCP/IP network, allowing users to view and **manipulate information** in multiple ways. Advanced **navigational tools** help users orient themselves and avoid becoming "lost in hyperspace".

[\[About\]](#) [\[Products\]](#) [\[Documentation\]](#) [\[FAQ\]](#) [\[Software\]](#) [\[News\]](#)

User: www-anonymous **Author:** gmesaric **Date:** 96/08/20 13:35:10

[Attributes](#)

[Parents](#)



User: www-anonymous



HYPERWAVE SERVER

KEY FEATURES

WORLD WIDE WEB

HyperWave Server software represents one of the most powerful WWW technologies currently available. Because it supports common network protocols and document formats, HyperWave Servers can be browsed and administrated with widely-used WWW browsers such as Netscape and Mosaic. Easy navigation is granted by hyperlinks and HyperWave's additional folder type: the collection.

- Integrated search engines
- Multilinguality
- Hyperlink consistency
- Meta-information
- Object-oriented database

INDUSTRY STANDARDS

Compatibility with industry standards is very important for WWW server technologies. By supporting standard network protocols such as HTTP/1.0, HyperWave Server provides unlimited connectivity to all kinds of HTTP clients and servers. The server provides support for Multi-purpose Internet Mail Extension (MIME) types and standard document formats such as HTML, GIF and JPEG. HyperWave Server also interacts easily with business applications using the Common Gateway Interface (CGI), the major standard gateway in the WWW.

ACCESS CONTROL

HyperWave Server stores all documents in an object-oriented database, it does not use the UNIX or Windows file systems. Providing its own access authorization system, HyperWave Server is much more secure than any other filesystem-based WWW server. It provides sophisticated access control to individual documents and collections using usernames, passwords, named groups and read/write/unlink rights. The server can be administrated remotely by all members of a special group "system". Billing and cashing - extremely important for Internet commerce - can be achieved using integrated accounting facilities.

REPLICATION

The HyperWave Interchange Format (HIF) lets users interchange interlinked multimedia webs between HyperWave Servers. Tools for importing and exporting collection trees to and from this format make it easy to replicate information on other servers. Applications reach from the mirroring of documentation or electronic journals to firewalls: HyperWave Servers let you create and modify public information on your side of the barrier and then replicate it to the "outside" server.

CONNECTIVITY

HyperWave Server provides connectivity to other applications and software modules via the Common Gateway Interface (CGI). This ensures interoperability with business applications, client/server databases, expert systems and special search engines. HyperWave Tools as a suite of command-line utilities provide a fast and robust interface to the server for batch programming and scripting. Finally the HyperWave SQL gateway ensures connectivity to common SQL databases: e.g. Oracle.

More attribute information.

[HOME](#) | [SEARCH](#) | [OPTIONS](#) | [IDENTIFY](#) | [EDIT](#) | [ANNOTATE](#) | [HELP](#)

HYPERWAVE AUTHOR

OVERVIEW

HyperWave Author is integrated hypermedia authoring software, specially designed to help you create and edit documents on HyperWave servers.

Special navigation tools such as collection browsers and hyperlink maps make the creation of high-quality webs easier than ever before. Integrated support for the HyperWave Server's search engines helps the user to find existing documents and reuse them. HyperWave Author will prevent you from getting "lost in hyperspace".

Distributed data management is the keyword for working with HyperWave software within large companies. The combination of HyperWave's access control features and HyperWave Author lets each department keep track of its information without having to set up its own server. This leads to a sharp decrease in software and administration costs.

- [KEY FEATURES](#)
- [TECHNICAL SPECIFICATIONS](#)
- Online documentation for [HyperWave Author for Windows \("Amadeus"\)](#)
- Online documentation for [HyperWave Author for UNIX \("Harmony"\)](#)

User: www-anonymous Author: gmesaric Date: 96/04/25 15:19:31

[Attributes](#)
[\[HyperWave Product Suite\]](#)

[HOME](#) | [SEARCH](#) | [OPTIONS](#) | [IDENTIFY](#) | [EDIT](#) | [ANNOTATE](#) | [HELP](#)

HYPERWAVE AUTHOR

KEY FEATURES

HYPERMEDIA AUTHORIZING

HyperWave Author is the ultimate interactive authoring tool for HyperWave Servers. Users can author remotely over network boundaries: the Internet or any other TCP/IP based network can be used. HyperWave author provides full support for the HG-CSP network protocol, special HyperWave Server features such as database and search facilities are seamlessly integrated into the interface.

OBJECT DATABASE

Object orientation is one of the key concepts of HyperWave. HyperWave Author provides full support for HyperWave Server's object-oriented database system, allowing easy insertion and editing of server-side objects. HyperWave Author for Windows additionally provides a local version of the database, letting users author web applications offline which they can later easily upload to a HyperWave Server.

VRML AND POSTSCRIPT

HyperWave Author software includes IICM's free VRML scene viewer and a viewer for PostScript documents. VRML is the standard 3D data format in the WWW. PostScript is the industry standard for electronic publishing. HyperWave Author provides integrated PostScript viewer software, including the facility for inserting hyperlinks in PostScript documents: annotations to non-HTML documents are possible because of HyperWave's link database approach.

ADVANCED NAVIGATION

Critics of the WWW often mention the so-called "lost in

hyperspace" syndrome. HyperWave Author provides advanced navigation concepts and demonstrates that there are solutions to this problem: tree-view collection browsers let you navigate easily through big web servers and dynamically generated hyperlink maps help you keep masses of interlinked information up to date.

DISTRIBUTED INFORMATION MANAGEMENT

HyperWave's authoring software provides the facility of distributed information management: every logical part of a company can have a virtual web server without having the overhead of setting up its own real web server. A company can have one corporate identity on the web, running a WWW service where every department of the corporation is responsible for its own part.

MULTILINGUAL DOCUMENTS

HyperWave Author supports easy creation and editing of multilingual web applications. HyperWave's support for multilingual document clusters is especially interesting if your company is located for example in Europe or Asia, or any other part of the world where more than one language is common. HyperWave Author's advanced navigational tools help you to get an overview of complicated multilingual webs.

User: www-anonymous **Author:** gmesaric **Date:** 96/04/17 10:28:14

[Attributes](#)

[[HyperWave Author](#)]

[HOME](#) | [SEARCH](#) | [OPTIONS](#) | [IDENTIFY](#) | [EDIT](#) | [ANNOTATE](#) | [HELP](#)



HYPERWAVE AUTHOR

TECHNICAL SPECIFICATIONS

HYPERWAVE AUTHOR FOR WINDOWS (AMADEUS)

- Compatible with industry standards
 - Supports HTML
 - Views common image formats such as GIF and JPEG
 - Integrated MPEG movie player
 - Comes with VRML (VRweb) and PostScript viewers
- 32-bit application (runs under Windows 3.1x using WIN32s)
- Efficient interactive connection to HyperWave Server (HG-CSP)
- Local object database for offline hypermedia authoring
- Sophisticated interface for HyperWave's integrated search engines
- Supports multilingual documents
- Full support for access control (identification, rights)
- Windows95 compliant TreeView for collection browsing

HYPERWAVE AUTHOR FOR UNIX (HARMONY)

- Compatible with industry standards
 - HTML edit API tool
 - Views standard image formats (GIF, JPEG, TIFF)
 - Integrated movie (MPEG) and audio (AVI, AU) players
 - Comes with VRML (VRweb) and PostScript viewers
- Advanced navigation tools (local map, 3D landscape)
- Interactive client-server protocol to HyperWave Servers
- Sophisticated interface to HyperWave's search engines
- Multilingual document management
- Support for access control (identification, rights)
- Point-and-click hyperlink creation
- Integrated communication facilities (talk/conference)

SUPPORTED PLATFORMS

HyperWave Author for Windows (Amadeus)

Vendor	Architecture	Operating System	Memory Requirements
Intel	486, Pentium	Windows95/NT	8 MB

HyperWave Author for UNIX (Harmony)

Vendor	Architecture	Operating System	Memory Requirements
Digital	Alpha, DEC RISC	OSF/1, Ultrix	64 MB
Hewlett-Packard	PA	HP-UX	64 MB
IBM	RS/6000	AIX	64 MB
Silicon Graphics	MIPS	IRIX	64 MB
Sun	SPARC	SunOS, Solaris	64 MB
Intel	486, Pentium	Linux, FreeBSD	32 MB

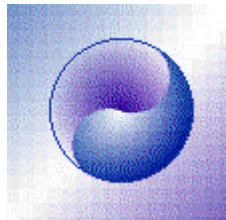
User: www-anonymous Author: gmesaric Date: 96/04/25 15:22:13

[Attributes](#)

[\[HyperWave Author\]](#)

[HOME](#) | [SEARCH](#) | [OPTIONS](#) | [IDENTIFY](#) | [EDIT](#) | [ANNOTATE](#) | [HELP](#)

Welcome to Harmony: The Hyperwave Administrator for Unix/X11



Harmony is the Unix/X11 client for [Hyperwave](#), the first second-generation, publicly available, networked hypermedia information system running over the Internet. Hyperwave integrates hyperlinking, hierarchical structuring, sophisticated search, and access control facilities into one single system, and is interoperable with other network information tools like Gopher and WWW.

See the [What's New](#) page for the latest news about Harmony.

Here you can find information about

- [Harmony Release 2.3](#)
- [where to get Harmony by anonymous FTP](#)
- Harmony's special features
 - [information structuring facilities](#)
 - [orientational aids](#)
 - [multilinguality](#)
 - Harmony's [document viewers](#)
- Installation guide
 - [installation](#)
 - [user configuration](#)
- User support

- the [Hyperwave mailing list](#)
- the [Harmony FAQ](#)
- further information

















Please direct any feedback (comments, suggestions, bug reports, etc.) concerning Harmony by electronic mail to:

harmony@iicm.tu-graz.ac.at

We are very interested in your feedback, even though we may not be able to respond personally to every piece of mail.

Author: kandrews, **Date:** 96/12/17 09:31:27, [Attributes](#)

HyperWave Author for UNIX ("Harmony")

-   [What's New with Harmony](#)
-   [Harmony Release Notes & Where to get Harmony by FTP](#)
-   [General Information About Harmony](#)
-   [Installation and user configuration of Harmony](#)
-   [Harmony User Support](#)
-   [Harmony- Frequently Asked Questions](#)
-   [The Harmony User Guide](#)
-   [Harmony Quick Reference](#)

User: www-anonymous **Author:** kandrews **Date:** 96/04/22 15:21:29

[Attributes](#)

[[HyperWave & Hyper-G](#)] [[HyperWave Online Documentation](#)]

[HOME](#) | [SEARCH](#) | [OPTIONS](#) | [IDENTIFY](#) | [EDIT](#) | [ANNOTATE](#) | [HELP](#)

Harmony's orientational aids

Harmony has many built-in features to discourage the phenomenon of "getting lost in hyperspace" while browsing large information spaces.



Local Map

Harmony's [Local Map](#) presents a dynamically generated graphical overview of the link relationships of a chosen document. Both incoming and outgoing hyperlinks are represented. Selecting an object toward the edge of the map and generating a new display offers a new means of associative browsing.

Location Feedback

When you select a document or collection in the Local Map, in the search result list, or follow a hyperlink, the location of the corresponding object in the collection hierarchy is **automatically** displayed in the collection browser, providing a powerful aid to orientation.



History

The [History Browser](#) offers a timeline of past interactive waypoints, including previous search panels.



3D Information Landscape

The [Information Landscape](#) is a three-dimensional graphical overview map of the collection structure. Users can "fly" over the hyperspace landscape looking for salient features, select interesting documents,

etc. The Mesa version (software only) is available on all platforms, the OpenGL version requires platform support and is available on SGI.

This feature requires platform support for OpenGL or Mesa and is currently available for SGI, DEC Alpha, Solaris, Linux and HP/UX machines.

User: www-anonymous **Author:** iicm **Date:** 96/12/17 14:00:37

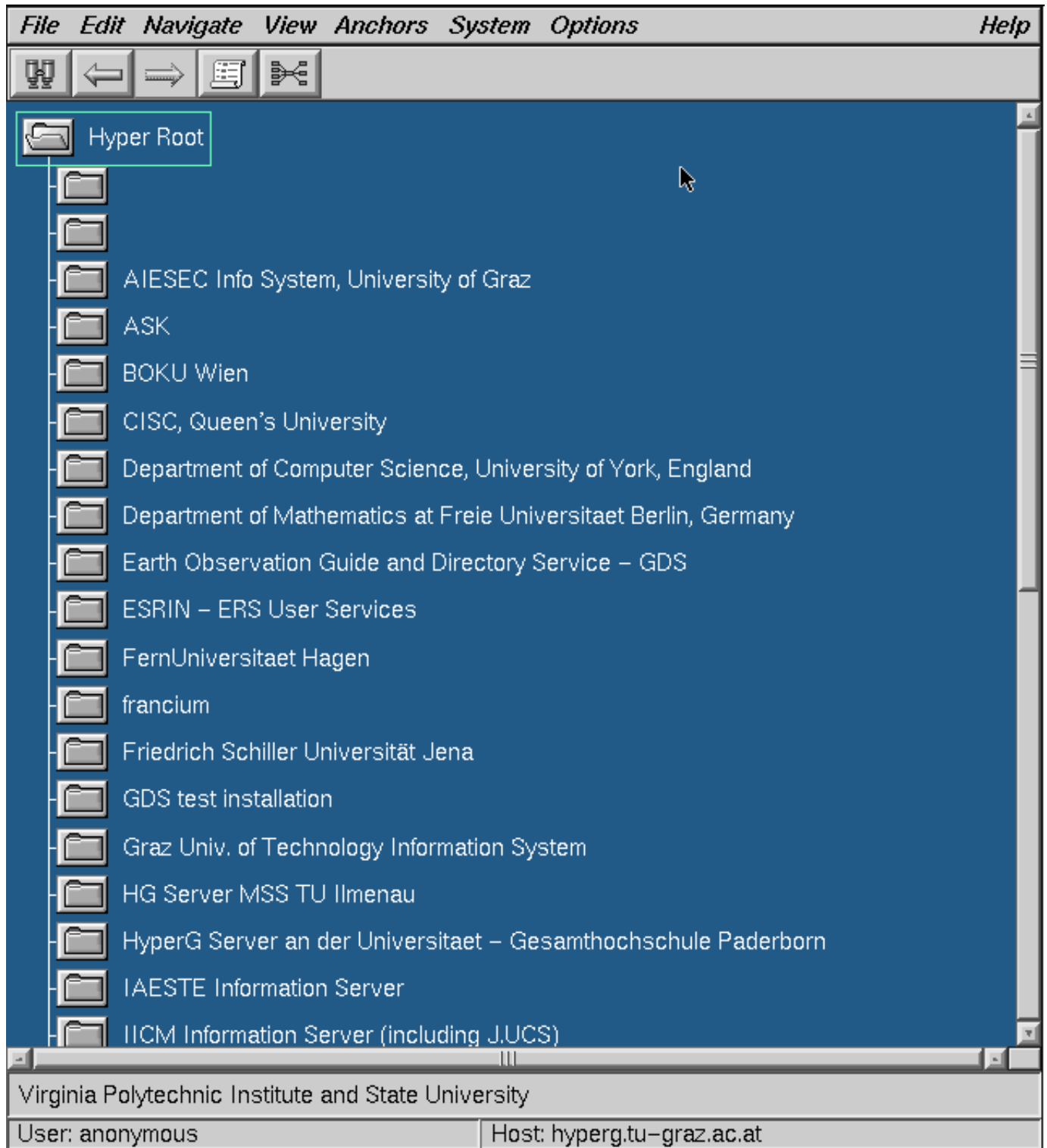
[Attributes](#)

[[Harmony's orientational aids](#)]

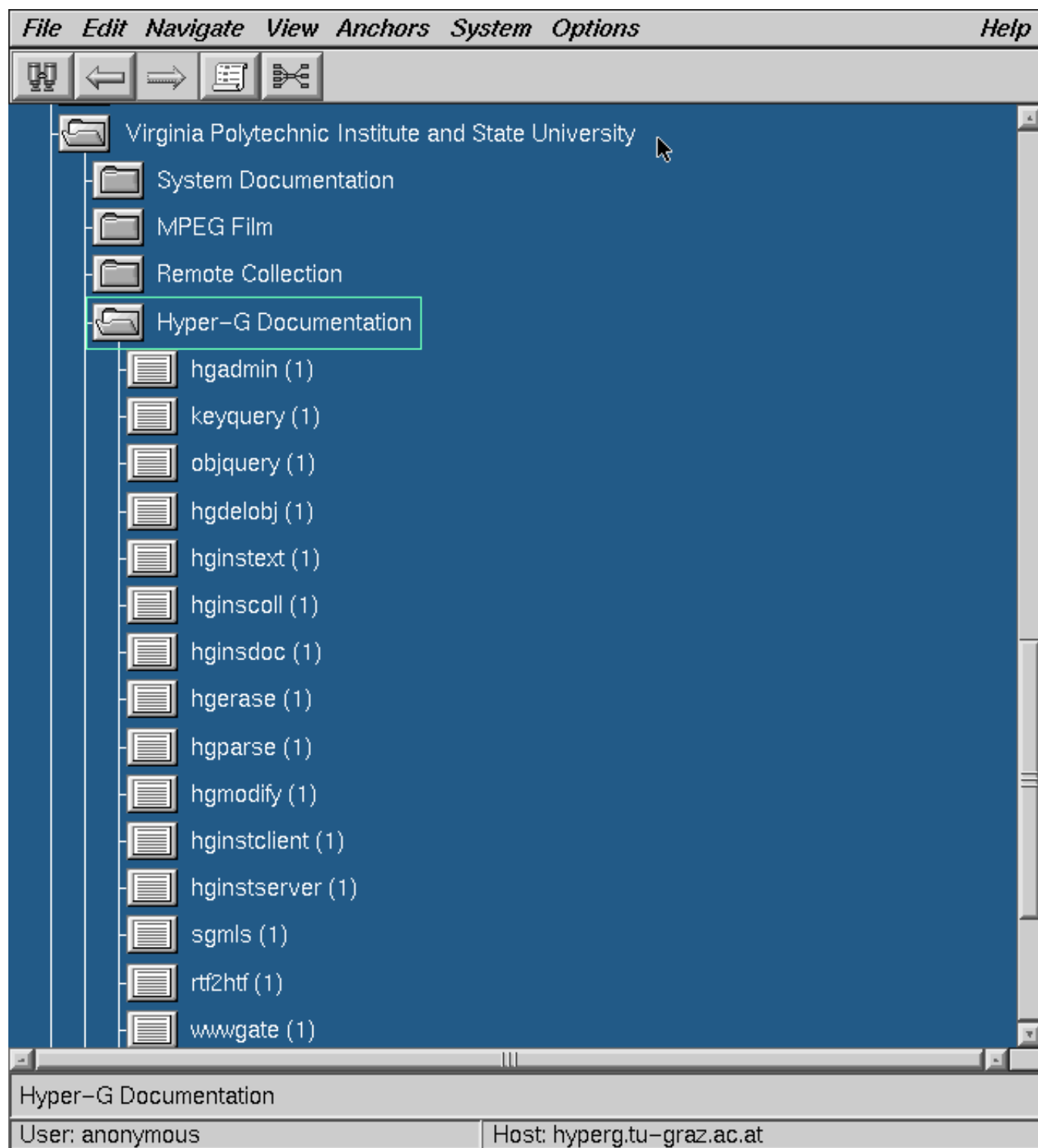
Hyper-G --- Harmony Illustrations

Illustrations of the use of Hyper-G with the Harmony (UNIX) client include:

- connection to the global root



- expansion of the collection of nodes accessible from the root to those at the Virginia Tech server



- viewing two text nodes and marking an anchor

File

Navigate

Anchors

View

Options

Help

Search

Anchors

hginscoll (1)

Name

hginscoll – insert a new collection

Synopsis

hginscoll [-h] [-i FCollId | -n FCollName] [-N CollName] [-c] [-A Author] [-C CDate] [-E EDate] [-O ODate] [-F][–T Title] [-R Rights] [-D Description] [-S SortOrder] [-L Language] [-r hgghost] [-d hgport]

Description

hginscoll builds a collection or cluster object and insert it into the Hyper–G database.

hginscoll (1)

Environment

HGAUTHOR: Author

HGRIGHTS: Rights

HGDESCRIPTION: Description

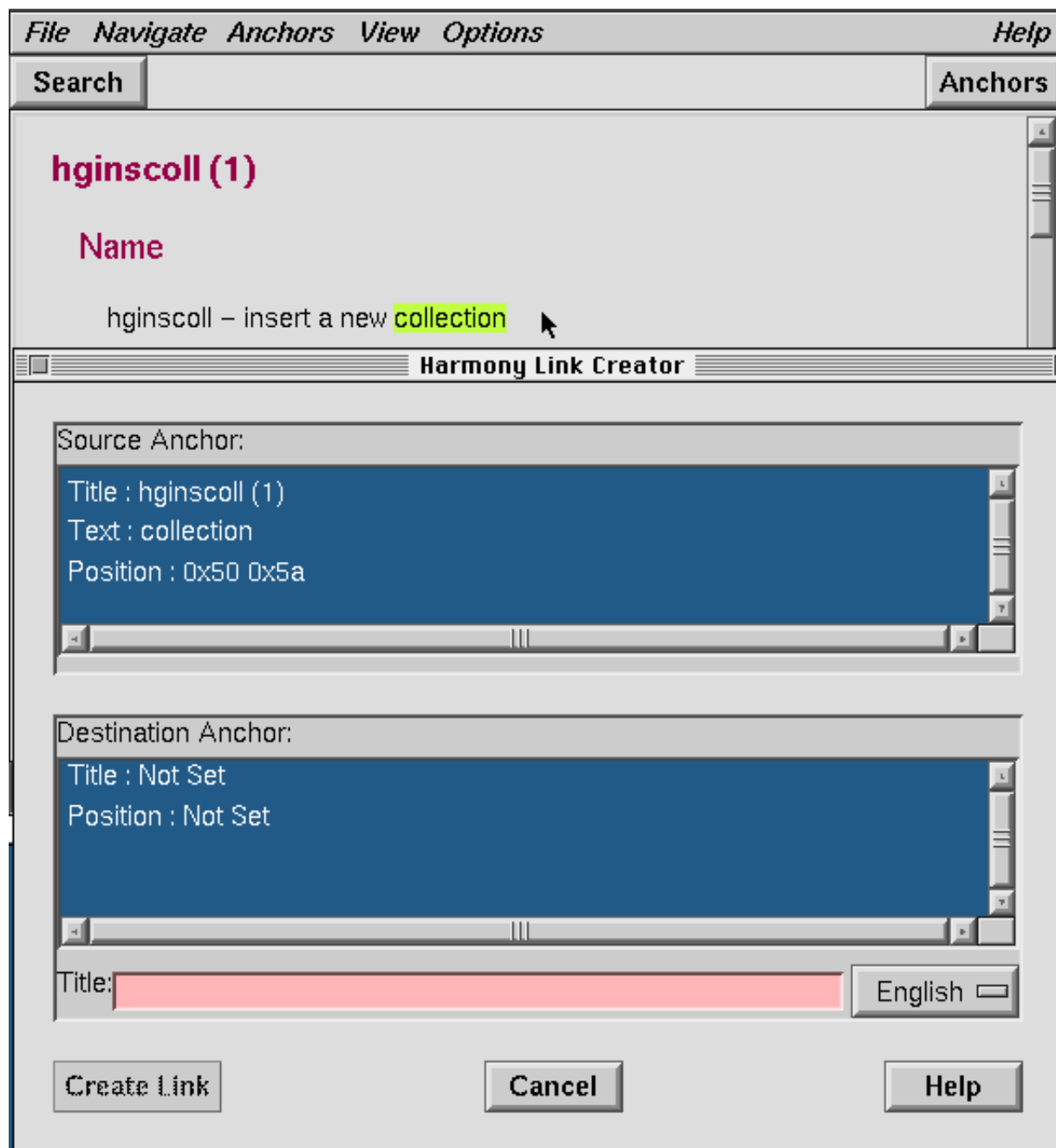
HGSORTORDER: SortOrder

HGFATHERCOLL: FCollName

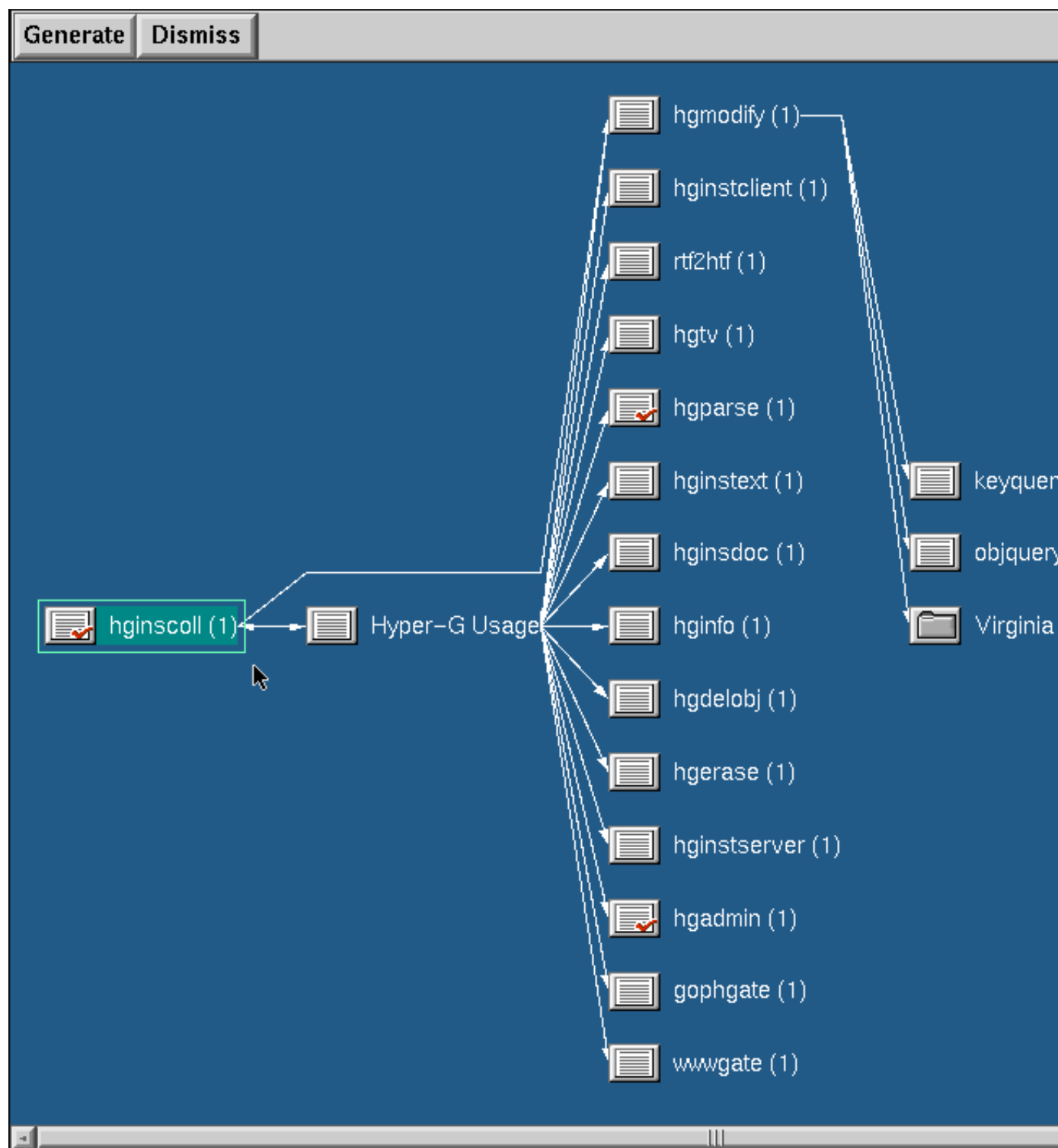
HGLANGUAGE: Language

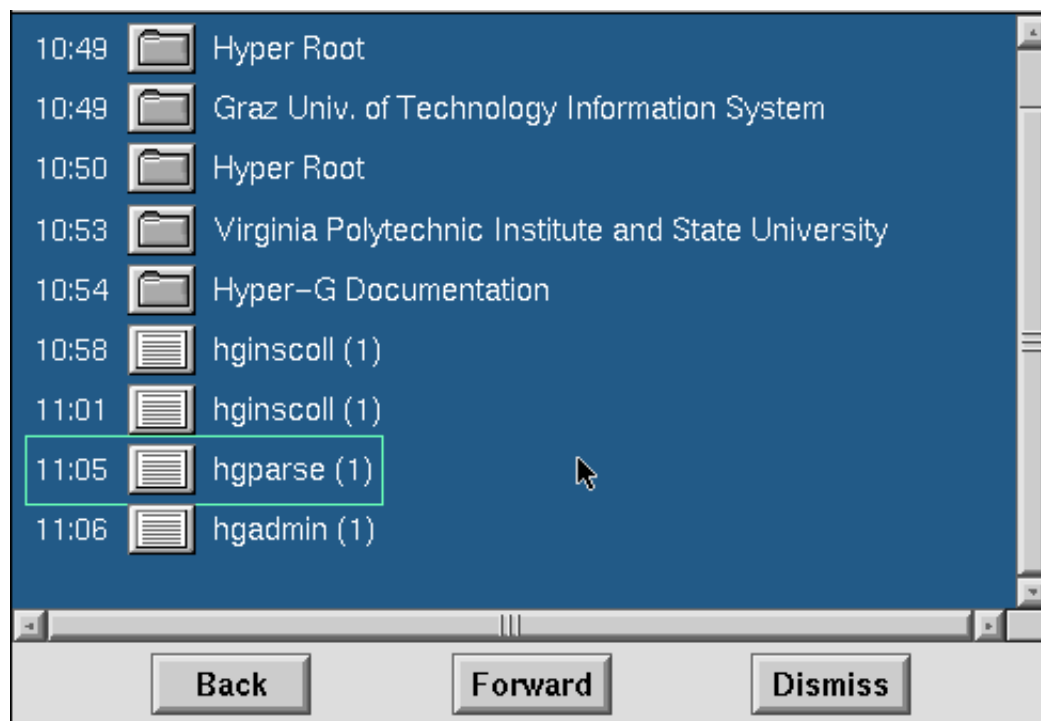
hginscoll (1)

- making a link



- viewing a local map





[HOME](#) | [SEARCH](#) | [OPTIONS](#) | [IDENTIFY](#) | [EDIT](#) | [ANNOTATE](#) | [HELP](#)

Information structuring facilities in Harmony

Hierarchical Browsing

Hyper-G servers use not only hyperlinks as a means of structuring information content but organize information into so-called **collections**, which are similar to directories, as well. All information on a Hyper-G server must be part of at least one collection, making it possible to access every document without the need for hyperlinks. Harmony's [Collection Browser](#) displays the hierarchical membership structure of Hyper-G data, like a graphical file browser and allows you to select and access objects that interest you.

Search

Harmony's [Search Dialog](#) supports both attribute (keyword, title, author, creation time, etc.) and content (full text) searches. The scope of searches is user-definable, ranging from individual collections to all collections on all Hyper-G servers worldwide. Search results are presented as a ranked list and can be used as the scope for further searches.

Hyperlinks

Harmony supports hyperlinks between arbitrary document types, including text, image, film, PostScript, and 3D scenes. Both source and destination anchors can be defined interactively.

User: www-anonymous **Author:** iicm **Date:** 96/12/17 14:01:04
[Attributes](#)



User: **www-anonymous**

Harmony's document viewers

Documents in Harmony are displayed by separate viewer processes in windows of their own:



Text Viewer

A generic SGML parser is used to display Hyper-G (HTF) and WWW (HTML) texts. Inline images in XBM, XPM, GIF, TIFF, and JPEG formats are supported.



Image Viewer

GIF, JPEG, TIFF, and PNG images are supported and may be zoomed, panned, etc. A special feature is live display -- when turned on, images are built up progressively on-screen as they are loaded. The autofit option automatically scales images to fit the current image viewer window.



Film Viewer

MPEG-1 video streams are supported. Options include live display while loading, double size display, alternative dithering methods, and gamma correction. After loading, selective portions of the film may be replayed, the frame rate altered, etc.



Audio Player

The Audio player is a graphical interface shell around whatever native audio command the system provides. The Audio Player supports both the Network Audio System (NAS) and local audio commands provided on your system. A full-featured, native Harmony Audio Player is under development.



PostScript Viewer

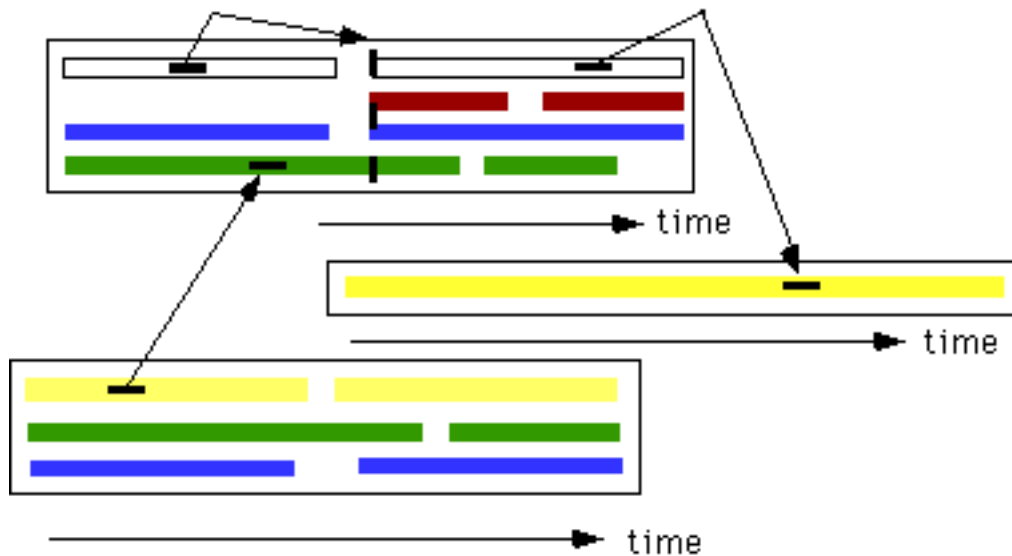
PostScript files can be displayed page by page, zoomed, printed, etc.



VRweb 3D Scene Viewer

3D model descriptions are displayed and can be manipulated or traversed in three dimensions. Hyperlinks are attached to objects in the model. The scene viewer is the Harmony version of VRweb, which supports models in VRML and SDF formats.





HYPERMEDIA COMPOSITES

[Adapted from HARD94 p. 53, Fig. 1c]

Models: Amsterdam - hypermedia

- Each of the 3 boxes represents a component or node that includes multimedia (time-based) information.
- When a component is presented, it *plays* for a period of time.
- The middle component is atomic, having only one media type. The black marker 2/3 of the way along is an anchor, which would allow one to *jump* directly to that point (in time / play).
- The other two components are composites, with several streams of multimedia information. The bottom one has three streams and six components, where components shown with the same color are played-back in sequence.
- The bottom component has one *from* anchor to a span in the first green-stream child-component of the top composite component.
- The top component is the most complex. About half-way through is a synchronization point, so one can jump there and all the streams will start then, at the same time. That point is reached from an anchor in the first white component. Other anchors point outside or are pointed to from outside.



Welcome to the **BEV HistoryBase**, a WWW History Page for the Blacksburg Electronic Village! Try out the BEV History Timeline to learn more about the history of our electronic community. For a non-graphical alternative, check out the Textual BEV History Timeline. Both contain the same information so feel free to browse either.



[[Main Timeline](#) | [Contribute](#) | [What's New?](#) | [Search](#)]

Message of the Day Listings

Blacksburg Telecommunications Advisory Committee Meeting Minutes

BEV Media Coverage Archive

BEV Group Home Pages

This project is supported by NSF Grant CDA-9424506. A copy of the grant proposal is online.

Last updated 27 October 1995 / schmidt@cs.vt.edu

[HistoryBase
Main Page](#)

[Contribute](#)

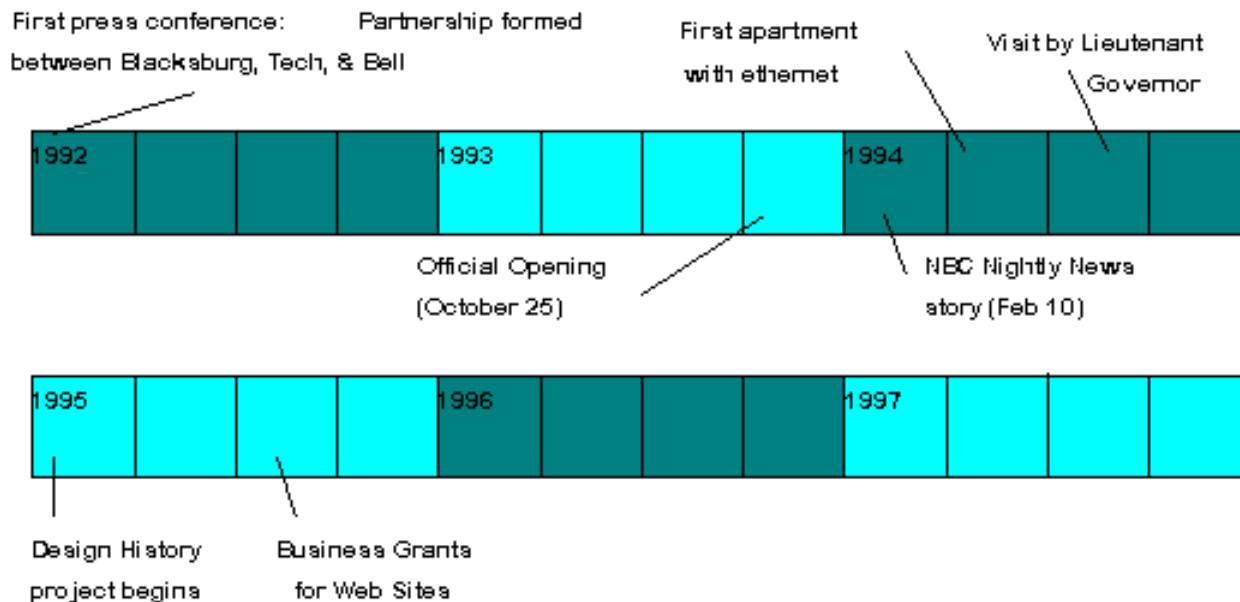
[What's
New?](#)

[Search](#)

[BEV
Homepage](#)

BEV HistoryBase: Main Timeline

Click in a box to see a more detailed history for that quarter



Click in a box to see a more detailed history for that quarter





[HistoryBase
Main Page](#)

[Contribute](#)

[What's
New?](#)

[Search](#)

[BEV
Homepage](#)







HistoryBase Main Page	 Main Timeline	 Prior Quarter	Next Quarter 	 Contribute	 What's New?	 Search
---	---	---	--	--	---	--

**Quick-Click
Timeline**

1 9 9 2 1 9 9 3 1 9 9 4 1 9 9 5 1 9 9 6 1 9 9 7

BEV HistoryBase: Jan-Mar, 1992

- January 18 1992 [Plan would change fiber of Blacksburg](#)
- January 21 1992 [Blacksburg: Model of computer future](#)
- January 21 1992 [Blacksburg may become 'electronic village'](#)
- January 21 1992 [Hi-tech may be the norm](#)
- January 21 1992 [Fiber optics may link Blacksburg](#)
- January 21 1992 [A look into the future](#)
- January 21 1992 [Project envisions Blacksburg as an 'electronic village'](#)
- January 27 1992 [Virginia Tech Launches Study for Fiber Optic Community Network](#)
- January 30 1992 [Electronic village proposed](#)
- January 31 1992 [Gut \(Comic Strip\)](#)
- February 1992 [Electronic Village: Technology showcase](#)
- February 06 1992 [Blacksburg Telecommunications Advisory Committee Minutes](#)
- February 11 1992 [Electronic village could make us lazy](#)
- February 13 1992 [In a Small Mountain Town, The 21st Century Is Calling](#)
- February 25 1992 [The Blacksburg Experiment](#)
- March 03 1992 [Blacksburg Telecommunications Advisory Committee Minutes](#)

HistoryBase Main Page	 Main Timeline	 Prior Quarter	Next Quarter 	 Contribute	 What's New?	 Search
---	---	---	--	--	---	--



TITLE: Blacksburg may become 'electronic village'

Contributor: KENNETH WILLIAM SCHMIDT JR (wschmidt@bev.net)

Submit Date: Sep 05 1995

Document Date: January 21 1992

Document Categories: Media Coverage: Newspaper

Author: Linda F. Jilk

Publication: *The News Messenger* (Montgomery County, VA)

Summary: This article contains the same information as all the other project announcement articles that appeared around this time. It describes the project at a very high level, gives a little technical information, describes some scenarios of how the BEV could be used, and presents a vision for the future with the project.

[Add an annotation to this document](#)





An Evolving Digital Library on Ancient Greece *and* Rome

Check out our preliminary [plans for a library of Shakespeare's sources and other Renaissance materials!](#)

[Copyright](#) | [General Policies of this Site](#)

Of the products of the soil, he allowed oil only to be sold abroad, but forbade the exportation of others; and if any did so export, the archon was to pronounce curses upon them, or else himself pay a hundred drachmas into the public treasury. His first table is the one which contains this law. One cannot, therefore, wholly disbelieve those who say that the exportation of figs also was anciently forbidden, and that the one who showed up, or pointed out such exporters, was called a "[sycophant](#)," or fig-shower.



Plutarch, [Solon](#), 24.1



Using the Perseus architecture catalog, Nicolas Stringos built a 1 cm:1.1 m scale model of the Parthenon on the Athenian Acropolis. Good work!

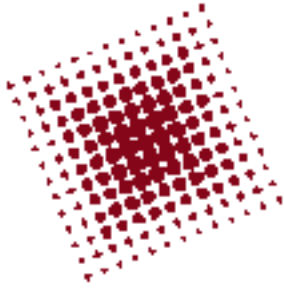
Take a look at [Nicolas' Parthenon model](#).

There is also a [text-only homepage](#).

The Perseus Project is supported by the  [Annenberg/CPB Project](#), the  [National Science Foundation](#), [Apple Computer](#) and the [National Endowment for the Arts](#), the [National Endowment for the Humanities](#), the Packard Humanities Institute, the [Getty Grant program](#), [Xerox Corporation](#), [Boston University](#), [Harvard University](#), and the [Fund to Improve Post-Secondary Education](#). Perseus is a non-profit enterprise, headquartered at the [Classics Department](#), [Tufts University](#).

Look at the [awards and reviews](#) the Perseus digital library has received.

Mail problems and suggestions to:
webmaster@perseus.tufts.edu



Corporation for National Research Initiatives

CS-TR Computer Science Technical Reports

- [An Introduction to the CS-TR Project](#), Robert E. Kahn, December 11, 1995
 - [Participants](#)
 - [Architecture of the Digital Library](#)
 - [Implementations](#)
 - [Contributed technology](#)
-

Participants

Each participant has provided on-line information about their work.

- [Carnegie Mellon University](#)
 - [Cornell University](#)
 - [University of California at Berkeley](#)
 - [Stanford University](#)
 - [Massachusetts Institute of Technology](#)
 - [CNRI](#)
-

Architecture of the Digital Library

Members of the CSTR project have been developing the basic architecture that must underlie a world wide digital library, where valuable information is stored. This work includes:

- An architecture for the digital library.
 - A [handle system](#) to maintain unique identifiers for objects in the Digital Library.
-

Implementations

Several public systems have been implemented with support from CSTR and are available for public use. (Some of these services are under development and subject to change at short notice.)

- [Dienst](#), a distributed search system for technical reports (Cornell)
- [Mercury](#), a centralized search system for technical reports (Carnegie Mellon)

Networked Computer Science Technical Reports Library

***NCSTRL** (pronounced "ancestral") is an international collection of computer science technical reports from CS departments and industrial and government research laboratories, made available for non-commercial and educational use. The NCSTRL collection is distributed among a set of interoperating servers operated by participating institutions. Read the official [NCSTRL press package](#) for a description of the background, goals, and organization of NCSTRL.*



Search the NCSTRL collection

- The **Fielded Search Form** allows you to perform a search on several fields of the bibliographic data, and/or to limit the search to specific institutions,
- **Or** enter one or several words into the box below to list all documents in our collection whose author, title, or abstract contain any search word:

- **Or** browse reports at any of the [participating institutions](#).

I want to join NCSTRL, tell me more

Read the [faq](#) for institutions interested in participating in the NCSTRL collection.

More Information

Find out [what's new with NCSTRL](#) or [browse a list of documents](#) related to NCSTRL.

NCSTRL at Cornell Computer Science. Send email to tech-reports@cs.cornell.edu.

[[Search](#) | [Home page](#)]

Simple Search Results

Search text:

hyperbase

Search Summary:

Organizations you selected are listed below by number of titles found.

- (1)[*Virginia Polytechnic Inst. and State University*](#)
- (1)[*Boston University*](#)

Search Results:

Virginia Polytechnic Inst. and State University

- [*A Query Language for Information Graphs.*](#) Sangita C. Betrabet, Edward A. Fox and Qi-Fan Chen. (TR-93-03)

Boston University

- [*Proceedings of the Workshop on Versioning in Hypertext Systems.*](#) David Durand, Anja Haake, David Hicks and Fabio Vitali. (95-001)

[[Search](#) | [Home page](#)]



NCSTR

*This server operates at Cornell University.
Send email to tech-reports@cs.cornell.edu*

Text Linking and Retrieval Experiments for Textbook Components

Gerard Salton, Chris Buckley and Zhongnan Zhao
TR90-1125
May 1990

Experiments are described designed to retrieve individual paragraphs of textbook material in answer to user-submitted queries. The retrieval strategies are based on the global comparison of paragraph texts, as well as on the local processing of text sentences. Furthermore, the retrieved items may be freely chosen, or may alternatively be restricted to certain areas in a clustered arrangement of book paragraphs. The retrieval results indicate that high retrieval values are obtainable for the more refined retrieval strategies, ranging between 0.70 and 0.80 in search precision.

How to view this document

- Display an **overview** of the document in one of the following formats.
 - [Overview of thumbnail pages](#)
 - [Structural overview](#)
- Display a **selected page** in one of the following formats (document has 14 pages).

raw OCR output		1	<input type="button" value="Display page"/>
hi-resolution tiff image			
inline gif image			

- Display the **whole** document in one of the following formats.
 - [OCR text](#) (produced by OCR, may have errors) 26029 bytes.
- [Print or download all or selected pages.](#)

You are granted permission for the non-commercial reproduction, distribution, display, and performance of this technical report in any format, BUT this permission is only for a period of 45 (forty-five) days from the most recent time that you verified that this technical report is still available from the Computer Science Department of Cornell University under terms that include this permission. All other rights are reserved by the author(s).

[[Search](#) | [Home page](#)]



NCSTR
This server operates at Cornell University.
Send email to tech-reports@cs.cornell.edu

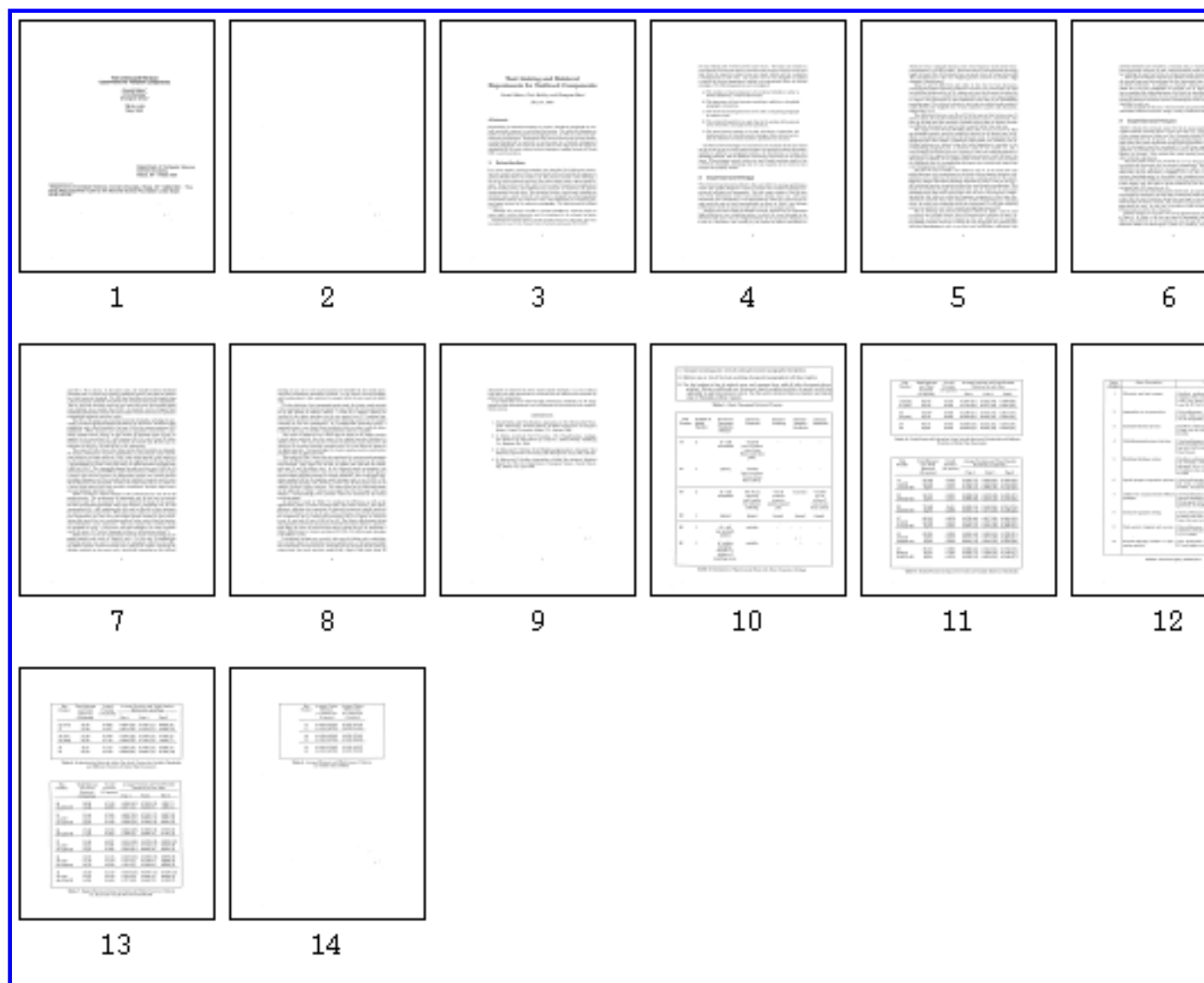
[[Prev section](#) | [Next section](#) | [Structure](#) | [Summary](#) | [Print/Download](#) | [Home page](#)]

Text Linking and Retrieval Experiments for Textbook Components

Gerard Salton, Chris Buckley and Zhongnan Zhao

May 1990

Section 1 of 1. Select a page.



[[Prev section](#) | [Next section](#) | [Structure](#) | [Summary](#) | [Print/Download](#) | [Home page](#)]



NCSTR

*This server operates at Cornell University.
Send email to tech-reports@cs.cornell.edu*

[[Prev page](#) | [Next page](#) | [Structure](#) | [Thumbnails](#) | [Summary](#) | [Print/Download](#) | [Home page](#)]

Click with the mouse to zoom in on a section of the page.

Text Linking and Retrieval Experiments for Textbook Components

Gerard Salton*
Chris Buckley*
Zhongnan Zhao*

TR 90-1125
May 1990

Department of Computer Sci
Cornell University
Ithaca, NY 14853-7501

*Department of Computer Science, Cornell University, Ithaca, NY 14853-7501. This study was supported in part by the National Science Foundation under Grant IRI-87-02735.

Digital Libraries for CS

Here are some pointers to Digital Libraries / bibliography servers related to CS.

[ACM Digital Library Collection at Virginia Tech](#)

Small test collection of CACM articles from those scanned in as part of the NSF-supported Envision project.

[ACM Graphics Bib. DB](#)

SIGGRAPH Online Bibliography Database

[ACM Computer Graphics Courseware Repository](#)

SIGGRAPH Computer Graphics Courseware Repository (ftp)

[ACM HCI Bib. DB](#)

interactions Bibliographies on Human-Computer Interaction

[BibNet Project](#) and [TeX Users Group](#) FTP bibliographies

bibliography collections from Nelson Beebe including HTML with extensive internal and external hypertext links. See examples: [IBM Systems Journal](#), [DEC Technical Journal](#). See [program to build these from BibTeX](#).

[CACM Collection \(1959-1979\) using Inquiry](#)

U. Mass. CIIR demo of Inquiry with CACM test collection

[Collection of Computer Science Bibliographies](#)

from Alf-Christian Achilles; updated monthly; 790 locally stored bibliographies; more than 530,000 references; 20,000 references contain URLs to an online version of the paper; more than 1600 links to other sites carrying bibliographic information; uses Glimpse

[Databases and Logic Programming \(mirror\)](#)

bibliography server by Michael Ley

[Hypertext Bibliography Project](#)

Hypertext Bibliography Project (Glimpse search of many publications)

[NCSTR](#)

Networked Computer Science Technical Report Library

[Table of Contents re LIS](#)

Table of Contents for JASIS, IPM, etc. - may be slow

[Univ. of Wales Cardiff CS Courseware](#)



ACM Digital Library Collection at Virginia Tech

This archive was created by the digital library project at Virginia Tech, in cooperation with ACM and IBM. You can search and browse the documents in the Computer Science literature.

Search the ACM Digital Library Collection

- The **Fielded Search Form** allows you to perform a search on several fields of the bibliographic data,
- Or simply enter one or several words into the box below to search author, title, and abstract:

- Or browse articles from the ACM Digital Library Collection.

Further Information

- Virginia Tech Digital Library project page.
- Virginia Tech Computing Center research department.
- Cruise a list of documents related to Networked Computer Science Technical Reports Library (NCSTRL) - Check out how to **participate** in NCSTRL, how to **download and install** the software, more technical details of the **Dienst** protocol and architecture (the technology behind NCSTRL), **links** to other sources of computer science technical reports, and more.

In case of problems send your comments to ACM digital library repository at Virginia Tech.



CS Dept. NSF-Supported Education Infrastructure Project / ei.cs.vt.edu

Welcome to the home page for the NSF-supported project "Interactive Learning with a Digital Library in Computer Science". We hope you find some of the results of our project useful! Please [send me comments and suggestions](#)! Thanks, Prof. E. A. Fox.

- [Courses](#) (over 40, with over 10K files)
- [Computer Science Education Innovation Workshop](#) June 15-21, 1997 **NEW!**
- QUIZIT [Software and thesis](#) **NEW!**;
Papers:
 - [QUIZIT: An interactive online quiz system for WWW-based instruction](#) - Tinoco, L. C., Fox, E. A., Ehrich, R. W, Fuks, H. In Proceedings of the VII Brazilian Symposium on Educational Technology. Belo Horizonte, Brazil, Nov. 1996.
 - Online Evaluation in WWW-based Courseware - Tinoco, L. C., Fox, E. A., Barnette, N. D. In Proceedings ACM SIGCSE'97, San Jose, Feb. 1997:
[paper in PDF](#), [presentation in PowerPoint](#)
- [Audio and Video Tutorials on Popular Tools and Systems](#)

- [Search ei.cs.vt.edu \(e.g., all courses\) using Harvest](#)
- [Project Overview](#)
- [ENVISION Project \(that led to development of the digital library\) Final Report](#)
- [SWAN \(algorithm visualization system\)](#)
- [Electronic Submissions of Student Programming Assignments used in CS3204](#)
- [Faculty Development Institute](#)
- [References](#)
 - [Project Overview \(for FIE'96, in PDF\)](#)
 - [Project Interim Report, Oct. 1996](#)
 - [Project Report for NSF EI PI Meeting, Nov. 1996](#)

This DEC Alpha, ei.cs.vt.edu, supports the Virginia Tech/Norfolk State University project "Interactive Learning with a Digital Library in Computer Science" that is funded in part by the National Science Foundation, 1993-1997.



CS Courses

Welcome to one of the largest (over 40 courses, over 10K files) repositories of Computer Science courseware! I hope you benefit and [send me comments and suggestions!](#)

Regards, Prof. E. A. Fox for

[Virginia Tech CS Dept.'s NSF Education Infrastructure Project](#)

- [CS1014: Numerical Computational Techniques](#)
- [CS1024: Computing For Business](#)
- [MaSc1044: Computer Science: A Liberal Arts Approach](#)
- [CS1044: Programming in C \(nonmajors\)](#)
- [CS1205: Operating System Tools I](#)
- [CS1344: Introduction to C Programming](#)
- [CS1206: Operating System Tools II](#)
- [CS1604: Computers and Networked Information](#)
- [CS1704: Introduction to Data Structures & Software Engineering](#)
- [CS2304: Self Study in a Programming System \(Java\)](#)
- [CS2304: Self Study in a Programming System \(UNIX\)](#)
- [CS2504: Introduction to Computer Organization](#)
- [CS2604: Data Structures and File Processing](#)
- [CS2704: Object-Oriented Software Design and Construction](#)
- [CS2964: Field Studies](#)
- [Honors 3004: Multimedia Technology and Projects](#)
- [UH3004: High-Performance Scientific Computing](#)

- [CS3204: Operating Systems](#)
 - [CS3304: Comparative Languages](#)
 - [CS/Math 3414: Numerical Methods](#)
 - [CS3604: Professionalism in Computing](#)
 - [CS3724: Introduction to Human-Computer Interaction](#)
 - [CS4104: Data and Algorithm Analysis](#)
 - [CS4114: Formal Languages](#)
 - [CS4124: Theory of Computation](#)
 - [CS4204: Computer Graphics](#)
 - [CS4214: Simulation and Modeling](#)
 - [CS4234: Parallel and Distributed Computing](#)
 - [CS4504: Computer Organization](#)
 - [CS4624: Multimedia, Hypertext and Information Access](#)
 - [CS4964: Field Studies](#)
 - [CS5014: Research Methods in Computer Science](#)
 - [CS5024: Models and Analysis](#)
 - [CS5034: Models of Computation](#)
 - [CS5114: Theory of Algorithms](#)
 - [CS5204: Operating Systems](#)
 - [CS5314: Concepts of Programming Languages](#)
 - [CS/EE5515: Computer Architecture](#)
 - [CS5604: Information Storage and Retrieval](#)
 - [CS5724: Models and Theories of HCI](#)
 - [CS6104: Symbolic Computation](#)
 - [CS6204: The World-Wide Web: Beyond the Basics](#)
 - [CS6404: Advanced Topics in Mathematical Software](#)
 - [CS6604: Interactive Accessibility](#)
-

[ACM Digital Library Collection at Virginia Tech](#)

Catalog Pages

- [Ugrad](#)
- [Grad](#)

Class Data Archives

- [EI Archives](#)
- [CS Department Archives](#)

[Searching All Courses and other Pages on ei.cs.vt.edu](#)

[Summary about Harvest collection from ei.cs.vt.edu](#)

[Usage Statistics](#)

All materials prepared for these [Dept. of Computer Science](#) courses are

Copyright 1995, 1996 [Virginia Tech](#)

Linking to or using these works for educational use is encouraged.

Commercial use of these works is strictly prohibited.

See also

- [CS listing for World Lecture Hall](#)
 - [NSF Computer Science Courseware Repository \(NSFCSCR\)](#)
 - [Computational Science Education Project](#)
-

Author: [Edward A. Fox](#)

Email: fox@vt.edu



CS5604 - Information Storage and Retrieval Fall 1996 - Table of Contents

- [Assignments](#)
- [Calendar](#)
- [Computers and Tools](#)
- [Course Format](#)
- [Course Notes / Overheads](#)
- [Department and Class Policies](#)
- [FAQ - Frequently Asked Questions](#)
- [Glossary \(in process\)](#)
- [Koofers \(old quizzes\)](#)
- [News / Announcements](#) (updated 961213@5am)
- [Photos of Class](#)
- **Projects:** [Initial Suggestions](#), [Groups](#), [Completed Projects](#)
- [Quizzes](#)
- [Readings and References](#)
- [Review](#)
- [Searching ei.cs.vt.edu Online with Harvest](#)
- [Status](#)
- [Syllabus](#)
- [Trips](#)
- **WWW Link Sets:** [Instructor's - CS4624: Multimedia, Hypertext and Information Access](#) - [WWW Virtual Library \(URLs organized by subject\)](#)

Pointers to Previous Years' Materials

- [Fall 1995](#)
 - [Debates](#)
 - [FAQ - Frequently Asked Questions](#)
 - [Summaries of Articles](#)
 - [Summaries of Class Sessions](#)
- [Fall 1994 and before](#)

[Usage Statistics](#)

Please send comments and suggestions to: fox@fox.cs.vt.edu

Calendar for CS5604

- 1. [DL](#): *Introduction, Digital Libraries*
FOX88d, SAMU91a, DONG87
8/27, 8/29
- 2. [IR](#): *Information Storage & Retrieval*
Ch1, SALT86a, Ch2
9/3, 9/5
- 3. [IF](#): *Inverted Files / Boolean Systems*
Ch3, Ch12, Ch15, SALT83d
9/10, 9/12
- 4. [IF](#) + [SS](#): *IF + String Searching*
Ch5
9/17, 9/19
- 5. [SS](#) + [RR](#): *SS + Ranking*
Ch10 + Ch14
9/24, 9/26
- 6. [RR](#): *Ranking / Relevance Feedback*
SALT75b, Ch11
10/1, 10/3 [at [Library](#)]
- 7. [CL](#): *Clustering*
Ch16
10/8, 10/10
- 8. [IN](#): *Indexing / Document Analysis*
Ch7, Ch8
10/15, 10/17
- 9. [SD](#): *SGML / Document Translation*
COOM87, MAMR87
10/22, 10/24
- 10. [HT](#): *Hypertext*
NIEL90a, HAAN92
10/29 [guest lecture by Dr. Jurgen Koenemann on feedback research], 10/31 [in Sun lab, McB 334, using KMS]
- 11. [HT](#) + [MM](#): *HT + Multimedia*
FOX91b, WALL91
11/5 [videotape *Interactive Digital Video*], 11/7 [demos in McB]

- 110]
- **12. MM: *Multimedia***
PHIL91a, GREE92
11/12, 11/14
- **13. KB: *Knowledge-Based Information Retrieval***
FURN87, LEBO88 (and, optionally, MALO87a)
11/19 [guest lecture by Robert France on KBIR, CODER,
MARIAN], 11/21 [guest lecture by Lucy Nowell and Robert
France on Envision]
- **Break** - 11/23 - Dec. 1
- **14. Catchup, Discussion, Project Presentations**
12/3, 12/5
- **15. Review**
12/10
- **16. Final Exam**
12/13 at 10:05am-12:05pm

Explanation

UNITS: There are 11 units in this course, each with a 2-letter ID that symbolizes the main topical area considered. Each unit will be covered in 1-2 weeks of class time, has a set of associated readings, and has some lab or homework exercises that must be completed. Either 10 or 15 points, depending on the time and difficulty of the unit, will be given when mastery is demonstrated by a quiz grade of at least 90% (or an equivalent sequence of successes through our new automated quiz system).

LINKS: The hypertext link for each unit above points to the related lecture notes.

WEEKS: There are 15 weeks in this course, with the 14th a break, and the last devoted to review, plus a final exam. The Calendar above gives details for each week. Its structure is according to the following template:

- **WEEK. ID:** *Unit Title(s)*

CS4624: Multimedia, Hypertext, Information Access Table of Contents (Spring 1997)

Please read: [News / Announcements](#)
(Most recent update is for [970508](#).)
READ NEWS!! [EXTRA POINTS!!](#) SIGN UP FOR [LISTSERV!!](#)

[Outline](#)

Student Information

- Home Pages for [Tuesday](#) and [Thursday](#)
- [Exercise Progress Table](#)
- [Tentative Grades](#)

Activities

- [Lab Sessions](#)
- [Lecture Notes](#)
- [Projects](#)
- Quizzes, Final: [Take](#), [Review Status](#) (ONLY FOR ENROLLED STUDENTS!) and [Pre-test](#), [Post-test](#)
- [Trips and Special Events](#)

Information Sources and Access Points

- [Class Inventory](#)
- [Computers and Tools](#)
- [Figures](#)
- [Figures with Captions](#)
- Glossary: [Local](#), [from IMA](#)

- [Index](#)
- [Koofers \(old quizzes, final\)](#)
- [Link Sets](#)
- [Listserv](#)
- [Tutorial on Multimedia \(1994\)](#)
- [Readings and References](#)
- [Searching All CS Class Notes Online with Harvest](#)

Syllabus

- [Calendar](#)
- [Department and Class Policies](#)
- [Instructor](#) and [GTA](#)
- [Syllabus Details: Grading, etc.](#)

[Spring 1995 Version](#)

[Spring 1996 Version](#)

[Comment Form --- Email Submission](#)

Copyright 1996, 1997 [Edward A. Fox](#)

CS4624 Outline

- [Pre-Test](#)
- IN: Introduction (1.5 wks), Labs:
 - Course Notes: [0](#), [1](#), [Exercises](#), Labs: [1](#), [2](#), [3](#), [Objectives](#), [Study Questions](#)
 - [Issues](#): Predictions, roles of publishers & electronic publishing, applications (e.g., multimedia mail, training, [video-on-demand](#)), hypertext/hypermedia
 - *Readings - textbook chapter*: [1](#)
 - *Systems/Demonstrations*: Author/Editor, KMS, MIME, Netscape, Storyspace
- AC: Application Construction (2.5 wks)
 - Course Notes: [2](#), [Exercises](#), [Objectives](#), [Study Questions](#)
 - [Issues](#): architectures / reference models (e.g., [IMA](#)), toolkits, authoring systems
 - *Readings - textbook chapters*: [7](#) (but not last part), [8](#); WWW: [Virtual Reality](#)
 - *Systems/Demonstrations*: AuthorWare, Director, ENVISION, Virtual Reality
- CR: Capture and Representation (3.5 wks)
 - Course Notes: [Audio](#), [3](#), [4](#), [5](#), [Exercises](#), Labs: [4](#), [Objectives](#)
 - [Issues](#): Universe of objects, human visual system, psychoacoustic modeling, digitization, quantization, conversion, metadata, cataloging, representing media (text, images, video, audio), CD-ROM
 - *Readings - textbook chapters*: [3](#), [4](#);
WWW: [CD-Video](#), [DVD](#); [File Formats \(Ms. State\)](#);
Extra Articles in CACM: Jan. 1994 CACM (Special Issue on Hypermedia),
Dec. 1991 CACM article on HyTime (Newcomb et al.),
April 1991 CACM (Special Issue on Digital Multimedia Systems) on JPEG (Wallace), MPEG (Le Gall), MPEG compression for CD-I (Sijstermans and van der Meer), px64 (Liou), MIME (Borenstein), and DVI Chips (Harney et al.)

- *Systems/Demonstrations/Exercises*: scanning, Illustrator, Photoshop, OCR, Photo-CD
- MC: Models and Compression (3 wks)
 - Course Notes: [6](#), [Exercises](#), Labs: [5](#), [Objectives](#), [Study Questions](#)
 - [Issues](#): Hypertext models (e.g., Dexter, Trellis, Amsterdam), standards (e.g., SGML, HyTime, MHEG, JPEG, MPEG), object classes, time, document models, vector quantization, DCT, entropy coding
 - *Readings - textbook chapters*: [2](#), [9](#) (part); Hyperbases: [ACM Hypertext on Hypertext](#), [Hypertext Compendium](#) - esp. [Dexter](#), [Trellis](#); book: [QuickTime](#); [page images](#) - esp. [Amsterdam](#), [IRIS](#) project, [Multimedia](#)
 - *Systems/Demonstrations/Exercises*: MPEG-player, DVI, QuickTime
- PI: Presentation and Interaction (2 wks)
 - Course Notes: [7](#), [Exercises](#), [Objectives](#), [Study Questions](#)
 - [Issues](#): rendering, browsing, navigation, searching & search engines, link services & engines, standards (e.g., PREMO), scripting
 - *Readings - textbook chapters*: [4](#) (review parts on display, output); [6](#) (last section); WWW: [Hyper-G](#), [ILDLCs](#), [OpenDoc](#) (and Bento), [QuickTime](#), [ScriptX](#)
 - *Systems/Demonstrations/Exercises*: [ENVISION](#), [HyperCard](#)/Toolbook, [Hyper-G](#), [MARIAN](#), [MHEG](#)
- NC: Networking and Communication (1.5 wks)
 - Course Notes: [8](#), [Exercises](#), Labs: [6](#), [Objectives](#), [Study Questions](#)
 - [Issues](#): distributed hypermedia databases, store & forward, Internet services, synchronization, quality of service, protocols, video conferencing, objects with their methods, multimedia servers
 - *Readings - textbook chapters*: [6](#) (not last section); [7](#) (last part); [9](#) (part); [10](#) (part); book: [Hodges+](#) WWW: [Information on Videoconferencing](#), [LBL Network Research Group Talks](#)
 - *Systems/Demonstrations/Exercises*: [CU-SeeMe](#), [MBone](#)
- [Post-Test](#) and [Review](#)
- [Quizes and Final](#)

CS4624 Projects

Completed Projects

- [Art and Architecture Slides](#)
- [Art History Slides](#)
- [ETD: Adobe Exchange](#)
- [ETD: conversion to PDF](#)
- [ETD: scanning](#)
- [Faculty Senate Expectations](#)
- [Humanities WWW pages](#)
- [Internet Audio](#) and [documentation](#)
- [MOOsburg](#)
- [Multimedia HCI Guidelines](#) and [6MB Director movie](#)
- Palace - use a [Palace client](#) to connect to 128.173.40.162 on port 9998
- [QBIC API](#)
- [QuickTime VR of campus](#)
- [Residence Hall Project](#)

[Project Overview](#)

[Project Grading](#)

[Project Ideas](#)

Chosen Project Ideas

See the student lists for the two labs to find assignments to projects (once determined), for [Tuesday](#) and [Thursday](#).

[Project Page for 1996](#)

Previous projects that may be of interest

1. [CAVE \(VR\)](#)
2. [Colds](#)
3. [Computer Literacy](#)
4. [Fashion](#)
5. [Libraries Virtual Tour](#)
6. [Tour Of Multimedia Labs on Campus](#)
7. [Honors 3004 - Multimedia Technology and Projects Guidelines, Old Projects](#)

[\[Home\]](#)

Copyright 1996, 1997 [Edward A. Fox](#)

Welcome to the On-line Virtual



You've reached the ticket window

Have you visited the Virtual Computer History Museum before?

- ☒ No
☐ Yes



Would you like to ?

© 1996, the [Virtual Computer History Museum group](#)
Comments may be emailed to museum@ei.cs.vt.edu



Welcome to the Virtual Computer History Museum

You've reached the Museum Lobby

There are several methods of viewing the historical exhibits contained within these electronic walls. How would you like to progress through the annals of computer history?

- ☒ A chronological look at the past (text-only)
- ☐ A chronological look at the past (graphical)
- ☐ Perform a keyword search
- ☐ Browse through our extensive gallery of images
- ☐ Take a tour of the exhibits
- ☐ Visit other related museums

Step Inside



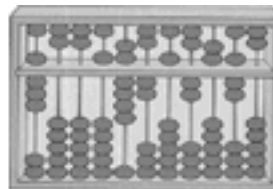
© 1996, the [Virtual Computer History Museum group](#)
Comments may be emailed to museum@ei.cs.vt.edu



Timeline of Events in Computer History

Click anywhere on the imagemap to display the specified timeframe in computer history

Early Years



1600's



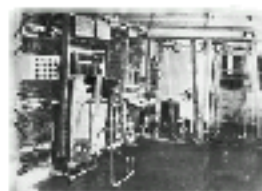
1800's



1900- WWII



WWII



post WWII

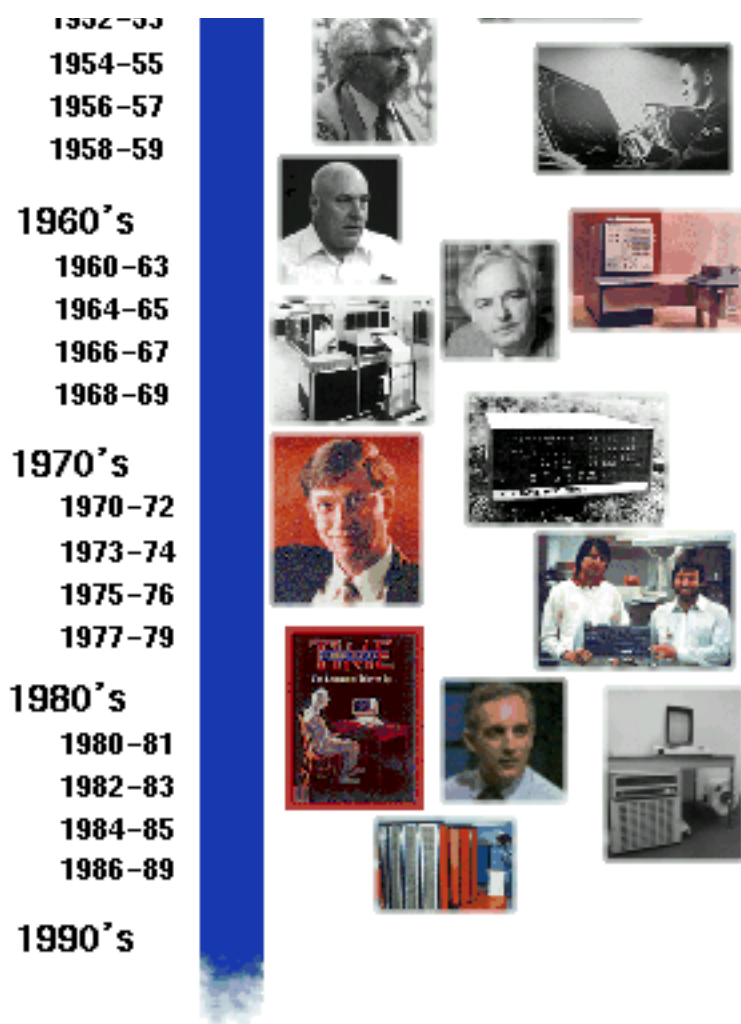


1950's

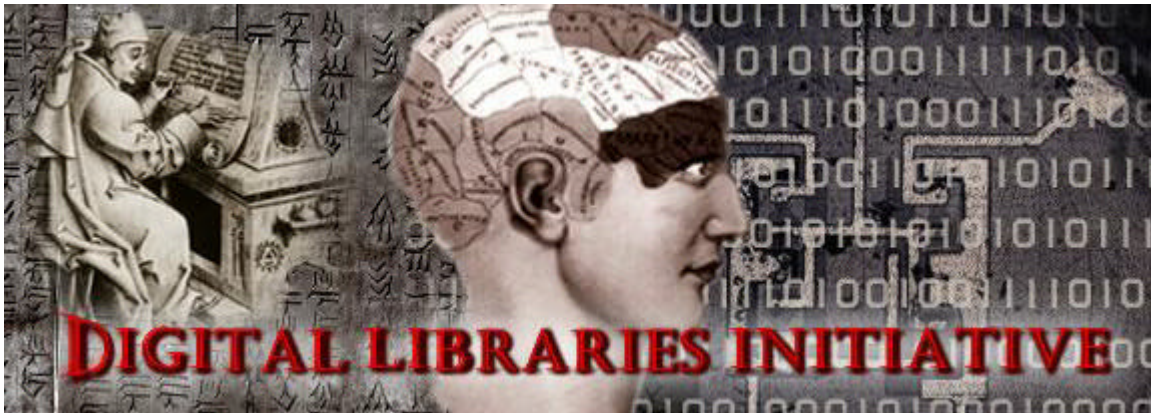
1950-51

1952-53





© 1996, the Virtual Computer History Museum group
Comments may be emailed to museum@core-dump.com



DIGITAL LIBRARIES INITIATIVE

Funded through a joint initiative of

[NSF/DARPA/NASA](#)

[University of
California at
Berkeley](#)

Environmental Planning
and Geographic
Information Systems

[University of
California at Santa
Barbara](#)

The Alexandria Project:
Spatially-referenced
Map Information

[Carnegie Mellon
University](#)

Informedia Digital Video

The Initiative's focus is to dramatically advance the means to collect, store, and organize information in digital forms, and make it available for searching, retrieval, and processing via communication networks -- all in user-friendly ways.

Digital Libraries basically store materials in electronic format and manipulate large collections of those materials effectively. Research into digital libraries is research into network information systems, concentrating on how to develop the necessary infrastructure to effectively mass-manipulate the information on the Net.. The key technological issues are how to search and display desired selections from and across large

Informedia Digital Video
Library

[University of Illinois
at
Urbana-Champaign](#)
Federating Repositories
of Scientific Literature

[University of
Michigan](#)
Intelligent Agents for
Information Location

[Stanford University](#)
Interoperation
Mechanisms Among
Heterogeneous
Services

DLI Project [Contacts](#)

[DLI Workshop
Series](#)

[DLI Publications](#)

collections.

Summaries of the six DLI projects from the May 1996, [Special Issue on Digital Libraries](#) in the Institute of Electrical and Electronics Engineers, IEEE Computer Magazine.

The magazine of digital library research, the [D-Lib Magazine](#), including the July/August 1996 issue [The DLI Testbeds: Today and Tomorrow](#).

Digital Library conference information, publications, related projects and resources to the DLI, [Digital Library Related Information and Resources](#).

[NSF Digital Libraries Contact](#)

National Synchronization for the Digital Library Initiative is being coordinated by the University of Illinois at Urbana-Champaign, and supported by a supplemental grant by the National Science Foundation.



Comments to [Susan Harum](#) at the University of Illinois DLI Project
10/31/96

Digital Library Related Information and Resources

The term digital library was amorphous before its status was elevated to buzzword. A digital library is more than a mere digitization of traditional library texts, as if that alone were an easy task. Research on digital libraries encompasses a range of intertwined technical, social and political issues. I have selected the following entries for their breadth, depth, ingenuity and availability of content online.

Table of Contents:

1. [The Digital Libraries Initiative \(DLI\)](#)
2. [Novel Digital Library Related Projects](#)
3. [Upcoming Digital Library Conferences](#)
4. [Previous Digital Library Conferences](#)
5. [Previous Digital Library Related Conferences with Online Proceedings](#)
6. [Full Text of Other Digital Library Related Publications](#)
7. [Other Digital Library Related Resources](#)
8. [Digital Library Funding, Coordination and Policy Organizations](#)
9. [Intellectual Property](#)
10. [Human Computer Interaction \(HCI\)](#)
11. [Computer Supported Cooperative Work \(CSCW\)](#)
12. [Networked Multimedia](#)

The Digital Libraries Initiative

The [Digital Libraries Initiative at the University of Illinois Urbana-Champaign](#) is one of six projects in the [joint initiative of the National Science Foundation \(NSF\), the Department of Defense Advanced Research Projects Agency \(DARPA\) and the National Aeronautics and Space Administration \(NASA\) for digital libraries](#).

These projects are in the process of developing the next generation of tools for information discovery, management, retrieval and analysis.

The other five DLI projects include: [Carnegie-Mellon University](#), [Stanford University](#), [University of California at Berkeley](#), [University of California at Santa Barbara](#) and [University of Michigan](#).

Novel Digital Library Related Projects

The [Interspace project](#) is the long term information infrastructure research portion of the NSF, DARPA and NASA Digital Library Initiative (DLI) at the University of Illinois Urbana-Champaign. The project seeks to unify disparate distributed information resources in one coherent model. The Interspace, is a collection of interlinked information spaces where each component space contains the knowledge of a community or a subject domain.

[Networked Computer Science Technical Reports Library](#) at Cornell University Department of Computer Science. A distributed technical report library developed by the ARPA-sponsored Computer Science Technical Report Project. "NCSTRL (pronounced "ancestral") is an international collection of computer science technical reports from CS departments and industrial and government research laboratories. The NCSTRL collection is distributed among a set of interoperating servers operated by participating institutions."

The Los Alamos National Laboratory(LANL) [Library Without Walls](#) is a broad based digital library project to make information available to researchers no matter where their desktops are located. The [LANL e-Print archive](#) "has already supplanted traditional research journals in some fields of physics. It is a formal mode of communication in which each entry is archived and indexed for retrieval at later times."

The [The CURIA Project's Thesaurus Linguarum Hiberni](#), "is a joint project of the Royal Irish Academy and the University College Cork to provide an interactive on-line searchable database archive of literary and historical materials in the various languages of early, mediæval and modern Ireland. The documents are being scanned from authoritative printed editions, or keyboarded from fresh manuscript transcriptions and encoded in SGML according to the

recommendations of the Text Encoding Initiative."

The [Perseus Project](#) at centered at the Department of Classics of Tufts University is a well known and respected collection which focuses upon the ancient Greek and Roman world. Perseus contains texts in Greek and in translation. The major authors of the classical period are represented, as well as some later authors from the fifth century B.C. Perseus also contains images of vases, sculptures and sculptural groups, coins, buildings, as well as color maps of Greece taken from satellite images, annotated with place names.

The [RYHINER-Project at the University Library of Berne](#) "consists of more than 15,000 maps, charts, plans and views from the 16th to the 18th century, covering the whole globe. Together with the 20,000 manuscript maps of the Public Records Office, the Canton of Berne owns not only a local, but a worldwide geographical memory. Work on this project includes conservation, microfilming and building up a generally accessible catalog."

[Project Bartleby](#) from Columbia University seeks to be the public library of the Internet. It reproduces classic literature for hypertext and maintains a strong emphasis on the quality and integrity of the text.

[Project Gutenberg](#) is the granddaddy of literary content on the Net. The goal of it's director and founder, Michael Hart, is no less than putting 10,000 works online by the year 2001. All works are in plain ASCII and in the public domain. In making the texts available to the lowest common denominator Project Gutenberg has the best chance to reach the most people and thus have the greatest impact.

Xerox has put together a collection called [Digital Libraries and Xerox](#) with papers discussing digital libraries and their research efforts. Xerox also has a number of interesting related projects including the [Digital Tradition Folk Song Database](#) which contains the words and music to thousands of folk songs. As well as the [Xerox PARC Map Viewer](#) uses public geographic data and renders sections of the world or the USA on the fly.

The [Visible Human Project](#) from National Library of Medicine (NLM) produced "a complete, anatomically detailed, three-dimensional representations of the male and female human body. The current phase of the project is collecting transverse CT, MRI and cryosection images of representative male and female cadavers at one millimeter intervals. The long-term goal of the Visible Human Project is to produce a system of knowledge structures that will transparently link visual knowledge forms to symbolic knowledge formats such as the names of body parts."

The [Digital Library Collection from the Library of Congress](#) is the beginnings of a National Digital Library which includes the American Memory project, Special American Collections at the Library of Congress and Country Studies.

The [Institute for Advanced Technology in the Humanities](#) has [research reports](#) about computing in the humanities at the University of Virginia, their online journal [Postmodern Culture](#), [technical reports](#) (with a section for SGML) and a forms based demonstration of the Institute's [Image Annotation Tool for Humanists](#).

The [IBM Digital Library](#) was an early commercial entry into the digital library arena. A major focus is on technical enforcement to copyright management.

Upcoming Digital Library Conferences

[ELVIRA4](#). The 4th UK Digital Libraries Conference (Electronic Library and Visual Information Research.) May 6-8 1997. Milton Keynes, UK.

[ADL '97](#). A Forum on Research and Technology Advances in Digital Libraries. May 7-9, 1997. Library of Congress, Washington, D.C.

[Digital Libraries '97](#). The Second ACM International Conference on Digital Libraries. July 24-26, 1997. Philadelphia, PA.

[International Summer School on the Digital Library 1997](#). August

10-22, 1997. Tilburg University, The Netherlands.

[AI in Digital Libraries](#). Part of the International Joint Conference on Artificial Intelligence Workshop Series. August 23-29, 1997, Nagoya, Japan.

[First European Conference on Research and Advanced Technology for Digital Libraries](#). September 1-3 1997, Pisa, Italy.

[IEEE Metadata 97](#). The Second IEEE Metadata Conference. September 16-17, 1997, Silver Spring, Maryland.

Previous Digital Library Conferences

This archives digital library conferences which have information online, but not full text of the proceedings.

[Visualizing Subject Access for 21st Century Information Resources](#) is the 34th Annual Clinic on Library Applications of Data Processing at the Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign. March 2-4, 1997. Urbana, IL.

[ACM DL'96](#). The First ACM International Conference on Digital Libraries. March 20-23, 1996, Bethesda, MD.

[ADL '96](#). Forum on Research and Technology Advances in Digital Libraries May 13-15, 1996, Washington, D.C.

[ELVIRA3](#). The UK Digital Libraries Conference. Third International Conference, Electronic Library and Visual Information Research. Hilton National Hotel, April 30-May 2 1996, Milton Keynes, UK.

[ADL '95](#). Research and Technology Advances in Digital Libraries. May 15-19, 1995. McClean Hilton at Tysons Corner, VA.

Previous Digital Library Related

Conferences with Online Proceedings

The following conferences and workshops have made all, or at least a substantial selection, of the full text of the proceedings available online.

[Libraries, People and Change: A Research Forum on Digital Libraries.](#)

The 38th Allerton Institute of the Graduate School of Library and Information Science University of Illinois at Urbana-Champaign. October 27-29, 1996. Allerton Park, Monticello, Illinois.

[Institute on Digital Library Development.](#) July 15-19 and July 29-August 2, 1996 Berkeley, California.

[IATUL 1996.](#) The International Association of Technological University Libraries. The overall theme of the conference will be "Networks, Networking and Implications for Digital Libraries." June 24-28, 1996 at the University of California, Irvine.

[OGDL II.](#) Organizing the Global Digital Library II and Naming Conventions May 21-22, 1996. Library of Congress, Washington, D.C.

[Digital Libraries '95](#) (DL'95). The Second International Conference on the Theory and Practice of Digital Libraries, held June 11-13, 1995. Austin, Texas.

[Digital Libraries '94](#) (DL '94). The First Annual Conference on the Theory and Practice of Digital Libraries June 19-21, 1994. College Station, Texas.

[ISDL'95.](#) International Symposium on Digital Libraries 1995. August 22-25, 1995. Tsukuba Science City, Ibaraki 305, Japan.

[IITA Digital Libraries Workshop.](#) Interoperability, Scaling and the Digital Libraries Research Agenda. May 18-19, 1995, Reston, Virginia.

[Social Aspects of Digital Libraries.](#) February 16-17, 1996, University of

California, Los Angeles.

[How we do user-centered design and evaluation of Digital Libraries: A methodological forum](#). The 37th Allerton Institute conference of the Graduate School of Library and Information Science University of Illinois at Urbana-Champaign. October 29-31 1995. Allerton Park, Monticello, Illinois.

[Information Gathering from Heterogeneous, Distributed Environments](#), the American Association for Artificial Intelligence (AAAI) Spring Symposium Series. March 27-29, 1995 Stanford University, Stanford, California.

[Building the Digital Library: Content Issues](#). Proceedings of the Library of Congress Network Advisory Committee. June 4-6, 1995. Library of Congress, Washington, D.C.

[Seminar on Cataloging Digital Documents](#) October 12-14, 1994 sponsored by the University of Virginia Library, Charlottesville and the Library of Congress.

[Adaptive hypertext and hypermedia workshop](#) held in conjunction with UM'94 (Fourth International Conference on User Modeling) Afternoon of Wednesday, August 17, 1994. Hyannis, Cape Cod, Massachusetts.

[TREC-4](#), the Text REtrieval Conference (TREC) is sponsored by the National Institute of Standards and Technology (NIST) and the Defense Advanced Research Projects Agency (DARPA). The articles are in Postscript.

[TREC-3](#), the Text REtrieval Conference (TREC). November 2-4 1994. Gaithersburg, Maryland. The articles are in Postscript.

[WWW5](#), the Fifth International World Wide Web Conference. May 6-10, 1996, at CNIT-Paris La Défense, France.

[WWW3](#), the Third International World-Wide Web Conference: Technology, Tools and Applications April 10-14, 1995, Darmstadt,

Germany.

[WWW2](#), the Second International World-Wide Web Conference: Mosaic and the Web. October 17-20, 1994, Chicago, IL.

[WWW1](#), the First International World-Wide Web Conference 25-27 May 1994, CERN, Geneva Switzerland.

Full Text of Other Digital Library Related Publications

These are pieces as well as collections that have been placed online in their full an unabbreviated form.

[D-Lib Magazine](#), an on-line, monthly magazine coordinated by CNRI and sponsored by DARPA on behalf of the IITA Working Group of the HPCC program, covers articles, news and commentary on advanced research and implementation projects in digital libraries.

[Buildings, books, and bytes](#) is the November 1996 by the Benton Foundation which reports on what library leaders and the public have to say about the future of libraries and communities in the digital age.

ERCIM - the European Research Consortium for Informatics and Mathematics has placed its [ERCIM News special theme on digital libraries](#) online. ERCIM News No.27 - October 1996.

The IEEE Computer Society's has placed the full text of related articles online for their [May 1996 theme issue of Computer on the US Digital Library Initiative](#).

[Solaris](#) is an annual review of research in information science and communications, including digital libraries from the Groupe interuniversitaire de recherche en sciences de l'information et de la communication (GIRSIC). The 1994, 1995 and 1996 are available in French with some English.

Many of the articles in the [SIGLINK Newsletter Special issue on Digital Libraries](#) are online. The articles are in a mix of HTML and Postscript. September, 1995 (Volume 4, Number 2).

An online edition of [Communications of the ACM - August 1995 Special Issue on Designing Hypermedia Applications](#).

The Association for Computing Machinery (ACM) has placed the full text of the April 1995 introduction online for the [Communications of the ACM issue on Digital Libraries](#).

The [Digital Library Source Book](#), 1993, edited by Edward Fox. The articles are in Postscript and PDF.

Other Digital Library Related Resources

These sites contain contain well rounded and or unique selections of information and resources about digital libraries.

The [Berkeley Digital Library SunSITE](#) is dedicated to gathering and publishing information about digital library projects and other digital content. It will also provide a platform for digital research and development as well as promote discussions on topics related to digital libraries, museums and archives.

The International Federation of Library Associations and Institutions or (IFLA) maintains a set of references for [digital libraries resources and projects](#), [metadata resources](#), [cataloging and indexing of electronic resources](#) and [interlibrary loan, document delivery and resource sharing information](#). IFLA also runs a number of mailing lists including the [DIGLIB](#) mailing list.

[New Horizons in Scholarly Communication](#) maintained by the Librarians Association of the University of California deals broadly with the use of new media in teaching and research, new publishing models and access issues. The section on access issues includes an

[introduction to the digital library](#).

An [annotated bibliography of digital library related sources](#) maintained Steven Ketchpel contains a wide array of annotated entries along with rankings for relevance and suggestions for intended audience .

The [Digital Libraries Resource Page](#) maintained by Karin L. Trgovac.

References on [Building Digital Libraries](#) from TexShare.

[WWW Library Resources - Discussion Lists](#) maintained by Randy D. Ralph contains descriptions and subscription information for many mailing lists related to digital libraries.

[Pointers to national and international library projects](#) from the BELNET User Forum Workgroup on Libraries.

Digital Library Funding, Coordination and Policy Organizations

The following organizations all provide explicit support or help contribute on a coordination or policy level to digital library related projects.

The [Corporation for National Research Initiatives](#) (CNRI) "is a non-profit organization dedicated to formulating, planning and carrying out national-level research initiatives on the use of network-based information technology." Many of their projects are digital library related.

The [Digital Library Technology](#) from the Information Sciences and Technology Branch Space Data and Computing Division NASA Goddard Space Flight Center. The DLT Project supports the development of new technologies to facilitate public access to NASA data via computer networks.

The National Coordination Office for High Performance Computing and Communications (HPCC) has made digital libraries a National Challenge Application in since 1995. "Blue Books" are annual reports presenting HPCC Program plans and accomplishments.

Digital Libraries References in the Blue Book

[HPCC: Advancing the Frontiers of Information Technology \(FY 1997 Blue Book\)](#)

[HPCC: Foundation for America's Information Future \(FY 1996 Blue Book\)](#)

[HPCC: Technology for the National Information Infrastructure \(FY 1995 Blue Book\)](#)

[National Digital Library Federation](#) (NDLF) is an organization constructed from fifteen of the nation's largest research libraries and archives.

Intellectual Property

Intellectual property rights and intellectual property rights management systems will be key issues and components of future digital libraries.

The [Intellectual Property Center](#) contains daily news with coverage of patents, copyright, trademark, Internet law, etc.

The [Intellectual Property Index](#) includes an index to resources on computer law, copyright, Internet law, multimedia law, trademarks and patents. The site also has a related discussion forum and is searchable.

The [Information Law Web](#) is a collection of links of people, place and things geared to helping people understanding their rights in terms of online information.

The [EFF Intellectual Property Online Archive](#) includes topics such as patents, trademarks and copyright contains a wide array of articles, legal documents and links to other resources in the area of intellectual

property.

The [Online Law Library](#) contains many high quality references to legal materials including [journals dealing with intellectual property issues](#).

The [WWW Multimedia Law](#) site producers and publishers of multimedia are oriented to legal liabilities faced on a number of platforms, not necessarily the Internet.

Human Computer Interaction (HCI)

The importance of user interfaces and human-computer interaction in general should not be underestimated with regard to digital libraries. Major advances in usability will come from innovation in the interfaces and not the underlying collections, databases or processing engines.

There are a number of good Human-Computer Interaction related sites on the web, one of these is the [Human-Computer Interaction Virtual Library](#) maintained by [Keith Instone](#).

Another is the [HCI resources](#) list maintained by [Mikael •Ericsson](#).

As well as an [index of HCI-related material in the Web](#) maintained by [Hans de Graaff](#).

The [ACM SIGCHI Home Page](#). SIGCHI is the ACM special interest group on Computer-Human Interaction. Conference proceedings from 1995 onward are available online.

Computer Supported Cooperative Work (CSCW)

Enhancement of collaborative and Cooperative forms of searching, communicating and creating are great advantages of the online medium and thus must be included into digital libraries.

A number of [CSCW](#) references including a [CSCW Bibliography](#), a large list of CSCW projects and products the [CSCW Yellow Pages](#) and [CSCW Related Links](#) have been compiled by [Michael Koch](#).

Contributors to the the USENet news group, [comp.groupware](#) have produced a number of FAQs which include the [comp.groupware FAQ hierarchy](#).

The [WWW Collaboration Projects](#) is a well rounded comprehensive site for applications on the Web that support some type of collaboration.

Networked Multimedia

Documents found in the digital library will consist of a broad range of media and formats of which text will be but a single type.

A broad [index to multimedia information sources](#) kept by Simon Gibbs.

The [MBONE home page](#) dedicated to the Multicast Backbone on the Internet is maintained by Vinay Kumar.

[WAXweb](#) is the hypermedia version of David Blair's feature-length independent film, "WAX or the discovery of television among the bees" (85:00, 1991). It combines one of the largest hypermedia narrative databases on the Internet with an authoring interface which allows users to collaboratively add to the story.

With the [Media on Demand](#) form you can request an immediate retransmission of sample MBONE seminars in both audio and video.

[Multimedia Integrated Conferencing for European Researchers](#) (MICE) is piloting Activities and the Conference Management and Multiplexing Centre. The site also contains the [MICE Multimedia Index](#) which has many links relevant to the subject of multimedia conferencing on the Internet.



Site for publications from all six Digital Libraries Initiative projects. Updated quarterly.

[University of California at Berkeley](#)
[University of California at Santa Barbara](#)
[Carnegie Mellon University](#)
[University of Illinois at Urbana-Champaign](#)
[University of Michigan](#)
[Stanford University](#)

University of California at Berkeley Digital Library Project

Belongie, S., Carson, C., Greenspan, H., Malik, J., [Recognition of Images in Large Databases Using Color and Texture](#), submitted to CVPR '97.

Buckland, M.K., Plaunt, C., [On the Construction of Selection Systems](#), Library Hi Tech, 12:4, 1994.

Carson, C., Ogle, V.E., [Storage and Retrieval of Feature Data for a Very Large Online Image Collection](#), IEEE Computer Society Bulletin of the Technical Committee on Data Engineering, Dec. 1996, Vol. 19

No. 4.

Fisher, D.E., [Topic Characterization of Full Length Texts Using Direct and Indirect Term Evidence](#), UC Berkeley Computer Science Technical Report UCB:CSD-94-809, May 1994.

Hearst, M.A., [Multi-Paragraph Segmentation of Expository Discourse](#), UC Berkeley Computer Science Technical Report UCB:CSD-94-790, January 1994.

Hearst, M.A., Plaunt, C., [Subtopic Structuring for Full-Length Document Access](#), Proceedings of the Sixteenth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, 1993, pp. 59-68.

Hearst, M.A., [TextTiling: A Quantitative Approach to Discourse Segmentation](#), UC Berkeley Computer Science Technical Report UCB:S2K-93-24, February 1994.

Kahn, R., Wilensky, R., [A Framework for Distributed Digital Object Services](#), Corporation for National Research Initiatives Technical report, May 1995.

Kopec, G.E., Lomelin, M., [Document Image Decoding Approach to Character Template Estimation](#), submitted to IEEE Trans. Pattern Analysis Machine Intelligence, Nov. 1995 (PostScript Version).

Larson, R.R., [Geographic Information Retrieval and Spatial Browsing](#), Presented at "GIS and Libraries: 32nd Annual Clinic on Library Applications of Data Processing" conference. University of Illinois at Urbana-Champaign, April 13, 1995. To appear in "GIS and Libraries: Patrons, Maps and Spatial Information", edited by Linda Smith and Myke Gluck.

Ogle, V.E., Stonebraker, M., [Chabot: Retrieval from a Relational Database of Images](#), IEEE Computer, Vol. 28, No. 9, September 1995.

Ogle, V.E., Wilensky, R., [Testbed Development for the Berkeley Digital Library Project](#), D-lib Magazine, July 1996, ISSN 1082-9873.

Palmer, D.A., Hearst, M.A., [Adaptive Sentence Boundary Disambiguation](#), UC Berkeley Computer Science Technical Report UCB:CSD-94-797, February 1994.

Plaunt, C., [A Workbench for Information Retrieval Experimentation](#), Association of Lisp Users Meeting and Workshop: Workshop Proceedings, 1995. Malvern, PA: Association of Lisp Users, 1995. Presentation prepared and given at the workshop in Cambridge, MA, August 1995.

Van House, N.A., Butler, M.H., Ogle, V.E., Schiff, L., [User-Centered Iterative Design for Digital Libraries: The Cypress Experience](#), D-lib Magazine, Feb 1996, ISSN 1082-9873.

Van House, N.A., Butler, M.H., Schiff, L., [The UC Berkeley Digital Environmental Library: Needs Assessment and Evaluation](#), October, 1995.

Van House, N.A., Butler, M.H., Schiff, L., [Needs Assessment and Evaluation of a Digital Environmental Library: the Berkeley Experience](#), October, 1995.

Van House, N.A., [User Needs Assessment and Evaluation for the UC Berkeley Electronic Environmental Library Project](#), Digital Libraries '95: The Second Annual Conference on the Theory and Practice of Digital Libraries, June 11-13, 1995 - Austin, Texas, USA.

Woodruff, A.G., Plaunt, C., [GIPSY: Automated Geographic Indexing of Text Documents](#), Journal of the American Society for Information Science, 45:9:645-655, 1994.

Buckland, M.K., What will collection developers do?, Information Technology and Libraries, September, 1995.

Forsyth, D.A., Malik, J., Fleck, M.M., Greenspan, H., Leung, T.,

Belongie, S., Carson, C., and Bregler, C., Finding Pictures of Objects in Large Collections of Images, ECCV '96 Workshop on Object Representation, Cambridge, April 1996.

Phelps, T.A., Wilensky, R., Multivalent Digital Documents in UC Berkeley's Digital Library Project, SIGLINK, Vol. 4, No. 2, September 1995.

Phelps, T.A., Wilensky, R., Toward Active, Extensible, Networked Documents: Multivalent Architecture and Applications, Proceedings of the 1st ACM International Conference on Digital Libraries, Bethesda, Maryland, March 20-23, 1996, pp. 100-108.

Phelps, T.A., Wilensky, R., Multivalent Documents: Inducing Structure and Behavior in Online Digital Documents , Best Paper Award - Digital Documents Track, Proceedings of the 29th Hawaii International Conference on System Science, Maui, Hawaii, January 3-6, 1996.

Wilensky, R., Toward Work-Centered Digital Information Services.

University of California at Santa Barbara Digital Library Project

Abdel-Ghaffar, K. El Abbadi, A., "Optimal Allocation of Two-dimensional Data," Proceedings of the 6th International Conference on Database Theory (ICDT'97), Delphi, Greece, January 1997, pp. 409-418.

Abreu, E., Lightstone, M., Mitra, S.K., Arakawa, K., "A New Efficient Approach for the Removal of Impulse Noise from Highly Corrupted Images," IEEE Transactions on Image Processing, v 5, June 1996 - to be published.

Andresen, D., Yang, T., Egecioglu, O., Ibarra, O., and Smith, T., "Scalability Issues for High Performance Digital Libraries on the World Wide Web." In Proceedings of ADL 1996, Forum on Research and Technology Advances in Digital Libraries, IEEE, Washington D.C., May 1996.

Andresen, D., Yang, T., Holmedahl, V., Ibarra, O., "SWEB: Towards a Scalable World Wide Web Server on Multicomputers," in Proceedings of IPPS '96 - 10th Inter. Parallel Processing Symposium, IEEE. Hawaii, April, 1996.

Andresen, D., Yang, T., Ibarra, O., "Towards a Scalable Distributed WWW Server on Networked Workstations" The Journal of Parallel and Distributed Computing (JP DC), 1996.

Andresen, D., Yang, T., Watson, D., Poulakidas, A., "Dynamic Processor Scheduling with Client Resources for Fast Multi-resolution WWW Image Browsing," Proceedings of the 11th International Parallel Processing Symposium (IPPS'97), Geneva, April, 1997.

Beard, K., Battenfield, B.P. "Detecting and Evaluating Errors by Graphical Methods." Chapter to appear in: Goodchild, M.F. Maguire, D. and Rhind, D. (Eds.) Geographic Information Systems: Principles and Prospects. London: Longman. (BB2, paper delivered December 1996; book forthcoming).

Beard, M.K., Smith, T.R., 1997. "A Framework for Meta-Information in Digital Libraries." In Managing Multimedia Data: Using Metadata to Integrate and Apply Digital Data. A. Sheth and W. Klaus (Eds.). McGraw Hill. (forthcoming).

Beard, M.K., Sharma, V., 1997. "Multidimensional Ranking in Digital Spatial Libraries." Special Issue of Metadata. Journal of Digital Libraries. (forthcoming).

Battenfield, B.P., "Delivering Maps to the Information Society: A Digital Library for Cartographic Data." Proceedings, 17th Conference of the International Cartographic Association, June 1997, Gavle, Sweden

(accepted; forthcoming).

Buttenfield, B.P., "GIS and Digital Libraries: Issues of Size and Scalability." In Smith, L.C. and Gluck, M. (eds.) GIS and Libraries: Patrons, Maps and Spatial Data. Champaign- Urbana: University of Illinois Press: 69-80.

Buttenfield, B.P., Goodchild, M.F., "The Alexandria Digital Library Project: Distributed Library Services for Spatially Referenced Data." Proceedings, GIS/LIS '96, Denver, Colorado, November 1996: 76-84.

Buttenfield, B.P., Kumler, M.P., "Tools for Browsing Environmental Data: The Alexandria Digital Library Interface." Proceedings Third International Conference on Integrating Geographic Information Systems and Environmental Modeling". Santa Fe, New Mexico, January 21-25, 1996 (no page numbers).

Buttenfield, B.P., Larsen, S., "Building a Digital Library for Earth System Science: The Alexandria Project." Invited Plenary Paper for Proceedings, Geology Society of America meetings, Denver, October 1996.

Buttenfield, B.P., Larsen, S., "Building a Digital Library for Earth System Science: The Alexandria Project." Proceedings Geological Society of America Meetings, Denver Colorado, November 1996 (paper delivered 15 January, proceedings forthcoming).

Buttenfield B.P., McLafferty, S., "Spatial Data in the Classroom: A Vision for Education. Proceedings, Conference on The Future of Spatial Data And Society. Washington, DC: National Academy of Sciences, Mapping Science Committee, and Federal Geographic Data Committee, 24-25 April, 1996.

Chandrasekaran, S., Manjunath, B.S., Wang, Y.F., Winkeler, J., Zhangk, H., "An Eigenspace update algorithm for image analysis," (CS TR, April 1996), accepted for publication in the journal CVGIP: Graphical models and image processing.

Chen, H., Chung, Y., Ramsey, M., Yang, C., "A Smart Itsy Bitsy Spider for the Web," Journal of the American Society for Information Science, Special Issue on AI Techniques for the Emerging Information Systems Applications, 1997, forthcoming.

Chen, H., Chung, Y., Ramsey, M., Yang, C., "An Intelligent Personal Spider (Agent) for Dynamic Internet/Intranet Searching," Decision Support Systems, 1997, forthcoming.

Chen, H., Chung, Y., Ramsey, M., Yang, C., Ma, P., Yen, J., "Intelligent Agents on the Internet," Proceedings of the 30th Annual Hawaii International Conference on System Sciences (HICSS-30), Maui, Hawaii, January 7-10, 1997.

Chen, H., Houston, A.L., Sewell, R.R., Schatz, B.R., "Internet Browsing and Searching: User Evaluation of Category Map and Concept Space Techniques," Journal of the American Society for Information Science, Special Issue on AI Techniques for the Emerging Information Systems Applications, 1997, forthcoming.

Chen, H., Martinez, J., Ng, T.D., Schatz, B.R., "Scientific Information Retrieval: An Experiment on the Worm Community System," Journal of the American Society for Information Science, Volume 48, Number 1, Pages 17-31, 1997.

Chen, H., Martinez, J., Kirchhoff, A., Ng, T.D., Schatz, B.R., "Alleviating Search Uncertainty Through Concept Associations: Automatic Indexing, Co-occurrence Analysis, and Parallel Computing," Journal of the American Society for Information Science, Special Issue on "Management of Imprecision and Uncertainty in Information Retrieval and Database Management Systems," 1997, forthcoming.

Chen, H., Ramsey, M., Li, H.P., "The Java Search Agent Workshop," submitted to IEEE Computer, Special Issue on Executable Content: Java and Its Applications, 1997.

Chen, H., Titkova, O., Orwig, R., Nunamaker, J.F., "Information

Visualization for Collaborative Computing," submitted to IEEE Computer, 1997.

Chen, H., Schatz, B.R., "Three Approaches to Controlling the Information Explosion on the Internet: Category Map, Concept Space, and Internet Agent," submitted to ACM Digital Library Conference, 1997.

Chen, H., Schatz, B.R., Houston, A.L., Sewell, R.R., Tolle, K.M., Doszkocs, T.E., Hubbard, S.M., "Challenges in Medical Information Retrieval: Evaluating Concept Space, Natural Language Parsers, and Category Map Techniques," submitted to ACM Digital Library Conference, 1997.

Chen, H., Schatz, B.R., Martinez, J., Ng, T., "Generating a Domain-specific Thesaurus Automatically: An Experiment on FlyBase," Information Processing and Management, 1997, forthcoming.

Chen, H., Shankaranarayanan, G., Iyer, A., She, L., "A Machine Learning Approach to Inductive Query by Examples: An Experiment Using Relevance Feedback, ID3, Genetic Algorithms, and Simulated Annealing," Journal of the American Society for Information Science, 1997, forthcoming.

Chen, H., Smith, T.R., Larsgaard, M.L., Hill, L.L., "A Geographic Knowledge Representation System (GKRS) for Multimedia Geospatial Retrieval and Analysis," International Journal of Digital Libraries, 1997, forthcoming.

Cheng, X., Dolin, R., Kothuri, R., Neary, M., Prabhakar, S., Wu, D., Agrawal, D., El Abbadi, A., Singh, A., Smith, T., Su, J., "Scalable Access Within the Context of Digital Libraries" to appear in Advanced Digital Libraries Forum, 97.

Dolin, R., Agrawal, D., El Abbadi, A., "Classifying Network Architectures for Locating Information Sources," Proceedings of the Fifth DASFAA Conference", April 1997.

Dolin, R., Agrawal, D., Dillon, L., El Abbadi, A., "Pharos: A Scalable Distributed Architecture for Locating Heterogeneous Information Sources," Computer Science Department, University of California, Santa Barbara, Technical Report Number TRCS96-05, 1996.

Frew, J., Freeston, M., Kemp, R., Simpson, J., Smith, T., Wells, A., Zheng, Q., "The Alexandria Digital Library Testbed", D-Lib Magazine, (July/August 1996).

Goodchild, M.F., "Scale in a digital geographic world," submitted to Geographical and Environmental Modeling, October 1996.

Houston, A.L., Chen, H., Schatz, B.R., Sewell, R.R., Tolle, K.M., Doszkocs, T.E., Hubbard, S.M., Ng, D.T., "Exploring the Use of Concept Space, Category Map Techniques, and Natural Language Parsers to Improve Medical Information Retrieval." Submitted to Decision Support Systems, Special Issue on Decision Support for Health Care in a New Information Age, 1997.

Ravi Kanth, K.V., Agrawal, D., El Abbadi, A., Singh, A., Smith, T.R., "Parallelizing Multidimensional Index Structures." IEEE Symposium on Parallel and Distributed Processing, New Orleans (SPDP 96): 376-383, October 1996.

Leitner, M., Buttenfield, B.P., "Cartographic Guidelines for Visualizing Attribute Accuracy." Proceedings, AUTO-CARTO 13, Seattle, Washington, April, 1997 (paper delivered 24 January, proceedings forthcoming).

Leitner, M., Buttenfield, B.P., "The Impact of Data Quality Displays on Spatial Decision Support." Proceedings, GIS/LIS '96, Denver, Colorado, November 1996: 882-894.

Ma, W.Y., Manjunath, B.S., "Edge Flow: A Framework of Boundary Detection and Image Segmentation" ECE Technical Report #97-02, University of California, Santa Barbara.

Ma, W.Y., Manjunath, B.S., "A pattern thesaurus for browsing large aerial photographs," ECE TR#96-10 (submitted for publication), June 1996.

Ma, W.Y., Manjunath, B.S., "Texture features and learning similarity," Proceedings IEEE Conf. on Computer Vision and Pattern Recognition, pp. 425-430, San Francisco, June 1996.

Ma, W.Y., Manjunath, B.S., "Tools for texture/color based search of images," To appear in SPIE Int. Conf. 3106, Human Vision and Electronic Imaging II, Feb. 1997.

Ma, W.Y., Manjunath, B.S., "A Texture Thesaurus for Browsing Large Aerial Photographs," (under revision for publication in the Journal of the American Society for Information Science, 1997, special issue on AI).

Manjunath, B.S., Ma, W.Y., "Texture features for browsing and retrieval of image data," IEEE Trans. Pattern Analysis and Machine Intelligence, August 1996 (special issue on digital libraries). Vol. 18, No. 8, pp. 837-842, August 1996.

Manjunath, B.S., Ma, W.Y., "Browsing large satellite and aerial photographs," Proceedings IEEE International Conference on Image Processing, Lausanne, Switzerland, September 1996 (invited paper), Vol. II, pp. 765-768.

Orwig, R.E., Chen, H., Nunamaker, J.F., "A Graphical, Self-Organizing Approach to Classifying Electronic Meeting Output," Journal of the American Society for Information Science, Volume 48, Number 3, March, 1997, forthcoming.

Poulakidas, A., Srinivasan, A., Egecioglu, O., Ibarra, O., Yang, T., "Experimental Studies on a Compact Storage Scheme for Wavelet-based Multiresolution Subregion Retrieval," in Proceedings of NASA 1996 Combined Industry, Space and Earth Science Data Compression Workshop, Utah, April 1996.

Prabhakar, S., Agrawal, D., El Abbadi, A., Singh, A., "Scheduling Tertiary I/O in Digital Libraries", submitted for publication.

Prabhakar, S., Agrawal, D., El Abbadi, A., Singh, A., "A Brief Survey of Tertiary Storage Systems and Research" to appear in ACM Symposium on Applied Computing (ACM SAC '97).

Prabhakar, S., Agrawal, D., El Abbadi, A., Singh, A., Smith, T., "Browsing and Placement of Multiresolution Images on Secondary Storage" Submitted for publication.

Skupin, A., Battenfield, B.P., "Spatial Metaphors for Display of Information Spaces." Proceedings, AUTO-CARTO 13, Seattle, Washington, April, 1997 (paper delivered 24 January, proceedings forthcoming).

Skupin, A., Battenfield, B.P., "Spatial Metaphors for Visualizing Very Large Data Archives." Proceedings, GIS/LIS '96, Denver, Colorado, November 1996: 607-617.

Smith, T.R., 1996. "The Alexandria Digital Library." D-Lib Magazine, March.

Smith, T.R., 1996. "The Meta-Information Environment of Digital Libraries." D-Lib Magazine, July.

Smith, T.R. et. al., 1996. "The Alexandria Digital Library: Overview and WWW Prototype," IEEE Computer, May, 1996, 54-60.

Smith, T.R., Geffner, S., Gottsegen, J., 1996. "A Framework for Catalogs and Metainformation in the Context of Digital Libraries," Proceedings of First IEEE Workshop on Metadata, Silver Spring, MD.

Strobel, N., Mitra, S.K., Manjunath, B.S., "Model-Based Detection and Correction of Corrupted Wavelet Coefficients," submitted to the 1997 IEEE International Symposium on Circuits and Systems, Hong Kong.

Strobel, N., Li, C.S., Castelli, V., "Texture-Based Image Segmentation

and MMAP for Digital Libraries," submitted to the 1997 IEEE International Conference on Image Processing, Santa Barbara.

Tolle, K.M., Chen, H., Chow, H., "Applying Neural Networks to Pharmacokinetic/Dynamic Data Sets," submitted to Decision Support Systems, Special Issue on Decision Support for Health Care in a New Information Age, 1997.

Tsou, M.H., Battenfield, B.P., "A Direct Manipulation Interface for Geographical Information Processing." Proceedings 6th International Symposium on Spatial Data Handling, Delft, The Netherlands, August 1996: 13B.37-13B.47.

Wu, D. Agrawal, D., El Abbadi, A., Singh, A., Smith, T.R., "Efficient Retrieval for Browsing Large Image Databases" Proceedings of the 5th International Conference on Information and Knowledge Management 96, pp. 11-18, Rockville, Maryland, Nov 12-16, 1996. (CIKM 96).

Wu, D., Agrawal, D., El Abbadi, A., Singh, A., "A Java-Based Architecture for Digital Library Storage" Proceedings of SPIE, Photonics East, Voice, Video, and Data Communications Boston, Massachusetts, November 23-24, 1996.

Carnegie Mellon University Digital Library Project

Christel, M., "[Addressing the Contents of Video in a Digital Library](#)." In the electronic proceedings of the ACM Workshop on Effective Abstractions in Multimedia: Layout, Presentation, and Interaction. Held in conjunction with ACM MM '95 on Nov. 4, 1995.

Christel, M., Wood, D., Stevens, S., [AMORE: The Advanced](#)

[Multimedia Organizer for Requirements Elicitation](#). Software Engineering Institute Technical Report CMU/SEI-93-TR-12. Pittsburgh, PA: Carnegie Mellon University, June, 1993.

Cox, B., Tygar, J.D., Sirbu, M., "[NetBill security and transaction protocol](#)" In Proceedings of the USENIX Workshop on Electronic Commerce. July, 1995(PostScript Version).

Hauptmann, A., Smith, M., "[Text, Speech, and Vision for Video Segmentation: The Informedia Project](#)," AAAI Fall 1995 Symposium on Computational Models for Integrating Language and Vision (PostScript Version).

Hauptmann, A.G., Witbrock, M.J., Christel, M.G., [News-on-Demand - An Application of Informedia Technology](#), D-LIB Magazine, September, 1995.

Hauptmann, A.G., Witbrock, M.J., Rudnicky, A.I., Reed, S., "[Speech for Multimedia Information Retrieval](#)" UIST-95 Proceedings of User Interface Software and Technology, Pittsburgh, PA, November, 1995 (PostScript Version).

Rowley, H., Baluja, S., Kanade, T., "[Human Face Detection in Visual Scenes](#)," Carnegie Mellon University, 1995. Computer Science Technical Report CMU-CS-95-158.

Smith, M., Christel, M., "[Automating the Creation of a Digital Video Library](#)," Proc. ACM Multimedia '95 (November 5-9, 1995), pp. 357-358.

Smith, M., Kanade, T., "[Video Skimming for Quick Browsing Based on Audio and Image Characterization](#)," Carnegie Mellon technical report CMU-CS-95-186, July 1995. Also submitted to PAMI Journal (Pattern Analysis and Machine Intelligence), 1995 (Compressed PostScript Version).

Wactlar, H., Kanade, T., Smith, M., Stevens, S., "[Intelligent Access to Digital Video: The Informedia Project](#)," IEEE Computer, Vol. 29, No.

5," May 1996.

Witbrock, M., [Learning Models of Speaker Variation](#). PhD Thesis. Carnegie Mellon University. May 1996 (PostScript Version).

Christel, M., Kanade, T., Mauldin, M., Reddy, R., Sirbu, M., Stevens, S., Wactlar, H., Informedia Digital Video Library. Communications of the ACM, 38(4):57-58 (1995).

Christel, M., Stevens, S., Wactlar, H., Informedia Digital Video Library. In: Proceedings of the ACM Multimedia '94 Conference. New York: ACM, October, 1994, pp. 480-481.

Christel, M., Stevens, S., Kanade, T., Mauldin, M., Reddy, R., Wactlar, H., "Techniques for the Creation and Exploration of Digital Video Libraries", to appear as Chapter 8 in Multimedia Tools and Applications (Volume 2), Borko Furht, editor. Boston, MA: Kluwer Academic Publishers, 1996.

Christel, M., The Role of Visual Fidelity in Computer-Based Instruction. Human-Computer Interaction, 9:183-223 (1994).

Christel, M., Wood, D., Stevens, S., Applying Multimedia Technology to Requirements Engineering. In: Proceedings of the Sixth Annual Software Technology Conference. Salt Lake City, UT: Software Technology Support Center, April, 1994 (CD-ROM proceedings).

Christel, M., Experiences with an Interactive Video Code Inspection Laboratory. In: Lecture Notes in Computer Science 640: Software Engineering Education SEI Conference Proceedings, C. Sledge, ed. Berlin: Springer-Verlag, October, 1992, pp. 395-411.

Hauptmann, A.G., "Speech Recognition in the Informedia Digital Video Library: Uses and Limitations", ICTAI-95 7th IEEE International Conference on Tools with AI, Washington, DC, November 6-8, 1995, (in press).

Hauptmann, A.G., Smith, M., "Video Segmentation in the Informedia

Project", IJCAI-95, Workshop on Intelligent Multimedia Information Retrieval. Montreal, Quebec Canada. August 1995.

Sirbu, M., Tygar, J. D., NetBill: An Internet Commerce System OPTimized for Network Delivered Services, Proceedings IEEE CompCon, March 1995.

Smith, M., Kanade, T., Video Skimming for Quick Browsing Based on Audio and Image Characterization, Carnegie Mellon technical report CMU-CS-95-186, July 1995.

Smith, M., Kanade, T., "Video Skimming for Quick Browsing based on Audio and Image Characterization," Submitted to the Special Issue of the IEEE Transactions on Pattern Analysis and Machine Intelligence on Digital Libraries: Representation and Retrieval, 1995.

Stevens, S., Christel, M., Wactlar, H., Informedia: Improving Access to Digital Video. interactions, 1(4):67-71 (1994).

Stevens, S., "New Models for Human-Computer Interaction in Digital Video Libraries" to appear in Bit.

Wactlar, H., Christel, M.; Stevens, S.; Kanade, T.; Mauldin, M.; Reddy, R.; Smith, M. "Technical Challenges for the Informedia Digital Video Library", to appear in proceedings of ISDL '95.

Wood, D., Christel, M., Stevens, S., A Multimedia Approach to Requirements Capture and Modeling. In: Proceedings of the First International Conference on Requirements Engineering. Los Alamitos, CA: IEEE Computer Society Press, April, 1994, pp. 53-56.

University of Illinois at Urbana-Champaign Digital

Library Project

Bishop, A.P., Star, S.L., Neumann, L., Ignacio, E., Sandusky, R.J., Schatz, B.R., [Building a university digital library: Understanding implications for academic institutions and their constituencies](#). In *Higher Education and the NII: From vision to reality*. Proceedings of the Monterey Conference, Washington, DC: Coalition for Networked Information, September 26-29, 1995.

Bishop, A.P., [Working towards an understanding of digital library use: A report on the user research efforts of the NSF/ARPA/NASA DLI projects](#). *D-LIB Magazine*, 1(3), October 1995.

[How We Do User-Centered Design and Analysis for Digital Libraries: A Methodological Forum](#) Collected papers from the 37th Allerton Institute, [Compiled by A. Bishop] October 29-31, 1995. [Compiled by A. Bishop].

Bishop, A.P., DLI Social Science Team Report, 11/15/95, [John and Jane Q. Engineer: What About Our Users?](#), November 15, 1995.

Bishop, A.P., Ignacio, E., Neumann, L., (working paper version) [Toward Functional Requirements for a Digital Library - Draft](#), December 1994.

Cain, A., Mcgrath, R.E., [Digital Commerce on the World Wide Web](#) NCSA Access, Summer 1995.

Harum, S., Mischo, W., Schatz, B.R., [Federating Repositories of Scientific Literature: An Update on the Digital Library Initiative at the University of Illinois at Urbana-Champaign](#), *D-Lib Magazine*, July/August 1996.

Ignacio, E., et al. [Usability Report](#), February 1996.

Johnson, E., Cochrane, P., [A Hypertextual Interface for a Searcher's Thesaurus](#). Editors: Shipman, F., Furuta, R. and David M. Levy, in

Digital Libraries '95, Austin, June 11-13, 1995.

Kanfer, A., [Digital Library Instrumentation Proposal](#) (working paper version) January 1995.

Kanfer, A., [Digital Library Instrumentation Outline - Draft](#) (working paper version) November 1994.

Krafft, D., [Proposed Infrastructure for a Shared Electronic Library of Documents for the Digital Library Initiative Awardees](#) (working paper version) January 1995.

Li, Y., Campbell, R., [A Dynamic Priority-based Scheduling Method in Distributed Systems](#), *Proc. of International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA'95)*, pp.177- 187, Georgia, November 1995 (PostScript Version).

Li, Y., Puvvada,V., Campbell, R., [Dynamic Retrieval of Remote Digital Objects](#), *Proc. of ACM Fourth International Conference on Information and Knowledge Management (CIKM'95)*, pp.182-187, Maryland, November 1995 (PostScript Version).

Li, Y., Chen, Z., Tan, S., Campbell, R., [Security Enhanced MPEG Player](#), to appear in *Proc. of IEEE First International Workshop on Multimedia Software Development (MMSD'96)*, Berlin, Germany, March 25-26, 1996 (PostScript Version).

McGrath, R.E., [UIUC DLI Project Scale-up: A Technical Evaluation](#).

McGrath, R.E., [Caching for Large Scale Systems: Lessons from the WWW](#) *D-Lib Magazine*, January 1996.

McGrath, R.E., Yeager, N. Cain, A. [World Wide Web servers at NCSA](#), *NCSA Access*, Spring 1995.

McGrath, R.E., [Performance of Several HTTP Demeons on an HP 735 Workstation](#), [Report] April 1995.

McGrath, R.E., [Comments on Haynes & Company CGI Benchmarks](#), (working paper version) November 1995.

Neumann, L.J., Ignacio, E.N., Sandusky, R.J., [Preliminary report on DLI Social Science Team data collection activities, 1994-1995](#). (working paper version), October 1995.

Powell, K. Schatz, B.R., [The Interspace System](#) (working paper version) November 1994.

Sandusky, R.J., [A Dynamic Repository for Organization Specific Information](#). [Report] December 1995.

Schatz, B.R. Chen, H., [Introduction to the Special Issue on Building Large-Scale Digital Libraries](#), *IEEE Computer*, 29(5) May 1996.

Schatz, B.R., Mischo, W., Cole, T., Hardin, J.H., Bishop, A.P., Chen, H., [Federating Repositories of Scientific Literature: The Illinois Digital Library Project](#), *IEEE Computer*, 29(5) May 1996

Schatz, B.R., [Information Analysis in the Net: The Interspace of the Twenty- First Century](#). refereed White Paper for *America in the Age of Information: A Forum on Federal Information and Communications R & D*, July 6-7, National Library of Medicine. sponsored by CIC (Committee on Information and Communications) reporting to the Science Advisor to the President of the United States.

Wedgeworth, R. Mischo, W., [Issues Affecting The Development of Digital Libraries in Science and Technology](#). Joint ICSU Press/UNESCO Expert Conference on ELECTRONIC PUBLISHING IN SCIENCE UNESCO, Paris, 19-23 February 1996.

Yeager, N., [Critique of Existing URN Implementations, when assessed against NCSA URN Requirements](#) (working paper version).

Bishop, A.P., Star, S.L., Social Informatics for Digital Library Infrastructure and Use. In: Martha Williams, ed. Annual Review of Information Science and Technology, vol. 30, in press, 1996.

Bishop, A.P., (Guest Editor). *SIGOIS Bulletin*, Special Issue on Digital Libraries, 16(2) 2-44, December 1995.

Bishop, A.P., Scholarly journals on the net: A reader's assessment. *Library Trends*, 43 (4), 544- 570, 1995.

Bishop, A.P., Squier, J., Artists on the Internet. In Kilnam Chon (Ed.). *INET '95 conference proceedings*, vol. II (1009-1018) Reston, VA: Internet Society, 1995.

Bishop, A.P, This little user went to market, this little user stayed home: What users, potential users, and nonusers can tell us. In *Challenging Marketplace Solutions to Problems in the Economics of Information: Proceedings*. Washington, DC: Association for Research Libraries, 1995.

Bowker, G., Star, S.L., How Things (Actor-Net)work, *Philosophia*, in press for early 1997.

Bowker, G. Star, S.L., Customized Software for Community Support versus Large Scale Infrastructure, Editors: G. M. Olson, J. B. Smith, and T. W. Malone, in *Coordination Theory and Collaboration Technology*, in press, 1995.

Chakrabarti, K., Mehrotra, S., Phantom Protection in R-trees, First Symposium on the Federated Lab on Interactive and Advanced Display, Army Research Labs, Aberdeen, MD, January 28-30, 1997.

Chakrabarti, K., Mehrotra, S., Concurrency Control in Multidimensional Access Methods, Technical Report, Department of Computer Science, University of Illinois at Urbana-Champaign, (to be submitted for publication), January 1997.

Chen, H., Martinez J., Ng, T.D, Schatz, B.R., A Concept Space Approach to Addressing the Vocabulary Problem in Scientific Information Retrieval: An Experiment on the Worm Community System, *Journal of the American Society for Information Science*,

48(1), 17-31, January 1997.

Chen, H., Vogel, D., Nunamaker, J.F., A Multi-Agent View of Strategic Planning Using Group Support Systems and Artificial Intelligence, *Group Decision and Negotiation*, 6(1), 35-58, January 1997.

Chen, H., Houston, A., Yen, J., Nunamaker, J.F., Intelligent Meeting Facilitation Agents: An Example on Group Systems, *IEEE Computer*, 29(8), August 1996.

Chen, H., Schatz, B.R., Ng, T.D., Martinez, J.P., Kirchhoff, A.J., Lin, C., A Parallel Computing Approach to Creating Engineering Concept Spaces for Semantic Retrieval: The Illinois Digital Library Initiative Project, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Special Section on Digital Libraries: Representation and Retrieval, 18(8) 771-782, August 1996.

Chen, H., Martinez, J., Ng, T.D., Schatz, B.R., A Concept Space Approach to Addressing the Vocabulary Problem in Scientific Information Retrieval: An Experiment on the Worm Community System, *Journal of the American Society for Information Science*, 47(8), August 1996.

Chen, H., Schuffels, C., Orwig, R., Internet Categorization and Search: A Machine Learning Approach, *Journal of Visual Communication and Image Representation*, Special Issue on Digital Libraries, 7(1) 88-102, March 1996.

Chen, H., Schatz, B.R., Ng, T.D., Yang, M., Breaking the Semantic Barrier: A Concept Space Experiment on the Convex Exemplar Parallel Supercomputer, submitted to *IEEE Parallel and Distributed Technology*, 1996.

Chen, H., Schatz, B.R., Ng, T.D., Yang, M., Large-scale Digital Library Analysis Using Supercomputers: An Experiment on CM-5, SGI Power Challenge, and Convex Exemplar, submitted to *IEEE Computer*, Special Issue on Applications for Shared-Memory Multiprocessors, 1996.

Chen H., Martinez, J., Kirchhoff, A., Ng, T.D., Schatz, B.R., Alleviating Search Uncertainty Through Concept Associations: Automatic Indexing, C-occurrence Analysis, and Parallel Computing, submitted to *Journal of the American Society for Information Science*, Special Issue on Management of Imprecision and Uncertainty in Information Retrieval and Database Management Systems, 1996.

Chen, H., Machine Learning for Information Retrieval: Neural Networks, Symbolic Learning, and Genetic Algorithms, *Journal of the American Society for Information Science*, 46(3): 194-216, 1995.

Chen, H., Martinez, J., Ng, T., Schatz, B.R., A Concept Space Approach to Addressing the Vocabulary Problem in Scientific Information Retrieval: An Experiment on the Worm Community System, *Journal of the American Society for Information Science*, 46(3): 175-193, 1995.

Chen, H., Ng, T., An Algorithmic Approach to Concept Exploration in a Large Knowledge Network (Automatic Thesaurus Consultation): Symbolic Branch-and-bound Search vs. Connectionist Hopfield Net Activation, *Journal of the American Society for Information Science*, Volume 46, Number 5, Pages 348-369, June 1995.

Chen, H., Schatz, B.R., A Path to Concept-based Information Access: From National Collaboratories to Digital Libraries, Editors: G. M. Olson, J. B. Smith, and T. W. Malone, in *Coordination Theory and Collaboration Technology*, 1995.

Chen, H., Schuffels, C., Orwig, R., Internet Categorization and Search: A Machine Learning Approach, *Journal of Visual Communication and Image Representation*, Special Issue on Digital Libraries, 1995.

Chen, H., Houston, A., Yen, J., Nunamaker, J., Intelligent Meeting Facilitation Agents: An Example on Group Systems, *IEEE Computer*, 1995.

Chen, H., She, L., Iyer, A., Shankaranarayanan, G., A Machine Learning Approach to Inductive Query by Examples: An Experiment Using Relevance Feedback, ID3, Genetic Algorithms, and Simulated Annealing, *IEEE Transactions on Knowledge and Data Engineering*, 1995.

Chen, H., Schatz, B.R., Lin, C., Concept Classification and Search on Internet Using Machine Learning and Parallel Computing Techniques, World Wide Web Conference-4 (WWW-4), Boston, December 11-13, 1995.

Chen, H., Schatz, B.R., Yim, T., Fye, D., Automatic Thesaurus Generation for an Electronic Community System, *Journal of the American Society for Information Science*, 46(3):175-193, April 1995.

Chen, H., Kim, J., GANNET: A Machine Learning Approach to Document Retrieval, *Journal of Management Information Systems*, 11(3): 9-43, Winter 1994-1995.

Chow, H., Chen, H., Ng, T., Myrdal, P., Yalkowsky, S.H., Using Backpropagation Networks for the Estimation of Aqueous Activity Coefficients of Aromatic Organic Compounds, *Journal of Chemical Information and Computer Sciences*, American Chemical Society, 35(4): 723-728, July/August 1995.

Cochrane, P., Johnson, E., Visual Dewey: DDC in a Hypertextual Browser for the Library User, *Advances in Knowledge Organization*, 5, 95-106, International Society for Knowledge Organization, Washington DC, 1996.

Cole, T., Digital Library Projects, *LITA Newsletter*, 16(2): 25-27, Spring 1995.

Cole, T., Kazmer, M., SGML As a Component of the Digital Library. *Library High Tech*, 13(4): 17-90, 1995.

Gross, B., Preserving and Securing the Electronic Record, Finding Common Ground Harvard College Library, Cambridge, MA, March

30-31, 1996.

Jacobson, F., Ignacio, E., Teaching Reflection: Information Seeking and Evaluation in a Digital Library Environment, To appear in *Library Trends*.

Johnson, E., Extending an Interactive Thesaurus by Dragging. *Siglink* Newsletter, Volume 4, Number 2, Pages 16-17, September 1995.

Kwan, T., McGrath, R.E., Reed, D.A., NCSA's World Wide Web Server: Design and Performance, *IEEE Computer*, 28(11):69-74, November 1995.

Laliberte, D., Hamilton, M., Experimental HTTP methods to support indexing and searching, June 1996.

Lagoze, C., McGrath, R.E., Overly, E., Yeager, N., A Design for Inter-Operable Secure Object Stores (ISOS). Cornell Computer Science Technical Report TR95-1558, 1995.

Lagoze, C., McGrath, R.E., Yeager, N., Overly, E., A Design for Inter-operable Secure Object Stores (ISOS), *ACM Int Conf on Digital Libraries*, 1995.

Li, Y., Tang, S., Sefika, M., Campbell, R., Liao, W., Experience with Scripting Kernel, Technical Report, Department of Computer Science, UIUC, November 1995.

Li, Y., Puvvada, V., Campbell, R., Index Management of the Digital Library Repository, Technical Report, Department of Computer Science, UIUC, May 1995.

Li, Y., The Design of Physical Support for Digital Libraries, Technical Report, Department of Computer Science, UIUC, December 1994.

Li, Y., Issues in Distributed File System and the Digital Library Project, Technical Report, Department of Computer Science, UIUC, December 1994.

Lin, C., Chen, H., An Automatic Indexing and Neural Network Approach to Concept Retrieval and Classification of Multilingual (Chinese- English) Documents, *IEEE Transactions on Systems, Man, and Cybernetics*, 26(1): 1-14, February 1996.

Neumann L., Star, S.L., Making Infrastructure: The Dream of a Common Language, Proceedings of the Participatory Design Conference 1996 (PDC '96), Cambridge, MA: Computer Professionals for Social Responsibility/ACM. Expanded version also submitted to *Journal of the American Society for Information Science*, 1996.

Orwig, R.E., Chen, H., Vogel, D., Nunamaker, J.F., A Multi-Agent View of Strategic Planning Using Group Support Systems and Artificial Intelligence, *Group Decision and Negotiation*, 6(1) 35-58, January 1997.

Orwig, R.E., Chen, H., Nunamaker, J.F., A Graphical, Self- Organizing Approach to Classifying Electronic Meeting Output, *Journal of the American Society for Information Science*, 1995.

Puvvada, V., Campbell, R., Inverse Mapping in the Handle Management System, Technical Report, Department of Computer Science, UIUC, October 1995.

Sandusky, R., Network Management as Cooperative Work: Implications for the Design of Integrated Network Management Systems, Submitted to The Fifth IFIP/IEEE International Symposium on Integrated Network Management to be held in San Diego, May 12-16, 1997.

Schatz, B.R., The Development Of The Information Sciences At UCSB, Report of an External Advisory Panel, February 1997.

Schatz, B.R., Information Retrieval in Digital Libraries: Bringing Search to the Net, *Science*, 275(5298), 327-334, cover story and lead article, January 17, 1997.

Schatz, B.R., Mischo, B., Cole, T., Hardin, J., Jackson, L., Bishop, A.P., Star, L.S., Cochrane, P., Chen, H., Digital Library Infrastructure for a University Engineering Community: Towards Search in the Net via Structure and Semantics, *IEEE Computer*, Special Issue on Building Large-scale Digital Libraries, May 1996.

Schatz, B.R., Chen, H., Building Large-Scale Digital Libraries, *IEEE Computer*, Special Issue on Building Large-scale Digital Libraries, 29(5) 22-27, May 1996.

Schatz, B.R., Johnson, E., Cochrane, P., Chen, H., Interactive Term Suggestion for Users of Digital Libraries: Using Subject Thesauri and Co-occurrence Lists for Information Retrieval, Proceedings from Digital Libraries '96: 1st ACM International Conference on Research and Development in Digital Libraries, Bethesda, MD, March 20-23, 1996.

Schatz, B.R., Building the Interspace: The Illinois Digital Library Project. *Communications of the ACM*, 38(4): 62-63, April 1995.

Schatz, B.R., Mischo, W., Cole, T., Hardin, J., Jackson, L., Bishop, A.P., Star, L., Cochrane, P., Chen, H., Digital Library Infrastructure for a University Engineering Community: Towards Search in the Net via Structure and Semantics, *IEEE Computer*, Special Issue on Building Large-Scale Digital Libraries, 1995.

Schatz, B.R., Johnson, E., Cochrane, P., Chen, H., Interactive Term Suggestion for Users of Digital Libraries: Using Subject Thesauri and Co-occurrence Lists for Information Retrieval, ACM Int Conf on Digital Library Conference, 1995.

Star, S.L., Grounded Classifications: Grounded Theory and Faceted Classifications, submitted to IFIP WG 8.2 Conference on Information Systems and Qualitative Research, June 1997.

Star, S.L., Working Together: Symbolic Interactionism, Activity Theory and Information Systems, in Yrjö Engeström and David Middleton, eds. *Communication and Cognition at Work*, Cambridge: Cambridge

University Press, in press, 1996.

Star, S.L., editor, *Cultures of Computing* (Sociological Review Monograph). Oxford: Basil Blackwell, 1995.

Star, S.L., Ruhleder, K., Steps toward an Ecology of Infrastructure, *Proceedings of CSCW 94*. New York: ACM Press, pp. 253-264.
Revised version accepted for Information Systems Research, special issue on Organizational Transformation, edited by JoAnne Yates and John Van Maanen.

Star, S.L., Bowker, G. Work and Infrastructure. (short contribution), *Communications of the ACM*, 38:41, September 1995.

Star, S.L.,. *Ecologies of Knowledge: Work and Politics in Science and Technology*. Editor. Albany: SUNY Press, 1995.

Star, S.L., Work and Practice in Social Studies of Science, Medicine and Technology, *Science, Technology and Human Values*, 20: 501-507, 1995.

Star, S.L., Infrastructure and Organizational Transformation: Classifying Nurses' Work (with Geoffrey Bowker and Stefan Timmermans) Pp. 344- 370 in *Information Technology and Changes in Organizational Work*, Orlikowski, W., Walsham, G., Jones, M., and DeGross, J., eds. . (Proceedings IFIP WG8.2 Conference, Cambridge, England.) London: Chapman and Hall.

Yeager, N.J., McGrath, R.E., Web Server Caching, *WebTechniques*, 1(2) 47-51, May 1996.

Yeager, N., McGrath, R.E., *Web Server Technology: An Advanced Guide for Information Providers*, Morgan Kauffman Publishers, in press, 1996.

Yen, J., Chen, H., Ma, P., Bui, T., An Issues Identifier for Online Financial Databases, *Proceedings of the International Society for Decision Support Systems Third International Conference, ISDSS'95*,

Hong Kong, June 22-23, 1995.

University of Michigan Digital Library Project

Atkins, D.E., Birmingham, W.P., Durfee, E.H., Glover, E., Mullen, T., Rundensteiner, E.A., Soloway, E., Vidal, J., Wallace, R., Wellman, M., [Toward Inquiry-Based Education Through Interacting Software Agents](#), Computer, May 1996, p. 69.

Birmingham, W.P. [An Agent-Based Architecture for Digital Libraries](#), D-Lib, July, 1995.

Birmingham, W.P., Durfee, E.H., Mullen, T., Wellman, M.P. [The Distributed Agent Architecture of the University of Michigan Digital Library](#). AAAI Spring Symposium on Information Gathering in Heterogeneous, Distributed Environments, Stanford, CA, AAAI Press (PostScript Version).

Kiskis, D. [UMDL Architecture Requirements Analysis](#). School of Information and Library Studies, University of Michigan, August, 1995. (Internal Report).

Mackie-Mason, J.K., White, K., [An Axiomatic Approach to Evaluating and Selecting Digital Payment Mechanisms](#). (PDF Version).

Mullen, T., Wellman, M.P., [A Simple Computational Market for Network Information Services](#). In First International Conference on Multiagent Systems, June 1995 (Compressed PostScript Version).

Price-Wilkin, J., [WWW-to-PAT Gateway: Exploiting an SGML-aware System Through the Web](#). The Public-Access Computer Systems Review 5, no. 7, 1994, 5-27.

Vidal, J., Durfee, E., [Task Planning Agents in the UMDL](#), Proceedings from CIKM '95: Workshop on Intelligent Information Agents, December 1-2, 1995.

Weise, J., [Visual Media Collections Browser for the Earth and Space Sciences](#).

Wellman, M.P., Durfee, E.H., Birmingham, W.P., [The Digital Library as Community of Information Agents](#), to appear in IEEE Expert, June 1996.

Atkins, D.E., Frank, R., Lougee, W., Willis, K., An Overview of Digital Library Initiatives at the University of Michigan, March 6, 1996.

Atkins, D.E., The Future of Libraries and Library Schools, Kanazawa Institute of Technology (KIT) International Roundtable for Library and Information Science, Japan, November 1995.

Atkins, D.E., A Library Educator's Perspective. Speech from the Research Library: The Day After Tomorrow, Association of Research Libraries, Proceedings of the 124th Annual Meeting, May 18-20, 1994, Austin, TX.

Birmingham, W.P., Drabenstott, K.M., Frost, C.O., Warner, A.J., Willis, K., The University of Michigan Digital Library: This is Not Your Father's Library. In Digital Library '94 Proceedings, June 1994, 53-60.

Bowen, W.G., The Foundation's Journal Storage Project (JSTOR), Report of the Andrew W. Mellon Foundation, 1994.

Cobbs, L., U-M project's helping America establish system for electronic libraries, The Ann Arbor News, November 6, 1994.

Crum, L., University of Michigan Digital Library Project. Communications of the ACM, Vol. 38, No. 4, April 1995, 63-64.

Crum, L., University of Michigan Digital Library Project, Final Report,

Volume I, Telepublishing Survey, TFPL Ltd., June 9, 1995.

D'Ambrosio, J., Birmingham, W.P., Preference-directed Design. AI in Engineering, Design, Analysis, and Manufacture, 1995 (to appear).

Darr, T.P., Birmingham, W.P., Automated Design for Concurrent Engineering. IEEE Expert, Vol. 9, No. 5, 1994, 35-42.

Durfee, E.H., Montgomery, T.A., Coordination as Distributed Search in a Hierarchical Behavior Space. IEEE Transactions on Systems, Man, and Cybernetics, Special Issue on Distributed Artificial Intelligence, Vol 21, No. 6, 1991, 1363-1378.

Gosling, W.A., Crist, M., Johnson, B., Lougee, W.P., Forrest Warner, B., Cooperative efforts in new methods of information delivery: the Michigan experience." Advances in Librarianship, vol. 19, 1995, p. 23-42.

Huebel-Drake, M., Finkel, L., Stern, E., Mouradian, M., Planning a Course for Success. Science Teacher, October 1995, Vol. 62, No. 7, p. 18-21.

Jackson, J., Enhancing access in a national digital library federation, New Edition, School of Information and Library Studies, University of Michigan, Spring 1995.

Jackson, J., Update: The U-M Digital Library Project, New Edition, School of Information and Library Studies, University of Michigan, Winter 1995.

Jacobson, R., Desktop libraries: University researchers strive to make vast networked collections a reality. The Chronicle of Higher Education, Vol. XLIL, No. 11, November 10, 1995.

Klein, A., Nasa, National Science Foundation give grants to "U." The Michigan Daily, October 11, 1994.

Lougee, W.P., Beyond access: new concepts, new tensions for

collection development in a digital environment." Collection Building, vol. 14, no. 3, 1995, p. 19-25.

Mackie-Mason, J.K., White, K., An Axiomatic Approach to Evaluating and Selecting Digital Payment Mechanisms.

Nesbit, J., Digital library project will receive \$8.5 million, The University Record, University of Michigan, Vol. 50, No. 7, October 17, 1994.

Nica, A., Rundensteiner, E.A., Uniform Structured Document Handling Using a Constraint-based Object Approach, Advances in Digital Libraries (ADL'95), A Forum on Research and Technology Advances in Digital Libraries, Virginia, May 1995 (PostScript Version).

Peters, P.E., Digital libraries are much more than digitized collections, Educom Review, July/August 1995.

Pool, R., Turning an Info-Glut Into a Library. Science, Vol. 226, No. 7, October 1994, 20-22.

Stix, G., The Speed of Write. Scientific American, December 1994, 106-111.

Wellman, M.P., Mullen, T., Market-Based Negotiation for Digital Library Services.

Woodford, J., The new "media union" on the North Campus: Building for the Future, Michigan Today, Vol. 27, No. 3, October, 1995.

Stanford University Digital Library Project

Balabanovic, M., Shoham, Y., [Combining Content-Based and Collaborative Recommendation](#) Communications of the ACM, 40(3), March, 1997 (PostScript Version).

Balabanovic, M., [An Adaptive Web Page Recommendation Service](#) (working paper version) In Proceedings of the First International Conference on Autonomous Agents, February, 1997.

Balabanovic, M., Shoham, Y., [Learning Information Retrieval Agents: Experiments with Automated Web Browsing](#) Proceedings of the AAAI Spring Symposium on Information Gathering from Heterogenous, Distributed Resources, March 1995. (Compressed PostScript Version).

Baldonado, M.Q., Chang, C.K., Gravano, L., Paepcke, A., [Metadata for Digital Libraries: Architecture and Design Rationale](#) (working paper version) Digital Libraries '97.

Baldonado, M.Q., Chang, C.K., Gravano, L., Paepcke, A., [The Stanford Digital Library Metadata Architecture](#) (working paper version) International Journal of Digital Libraries, 1(2), February, 1997.

Brin, S., Davis, J., Garcia-Molina, H., [Copy Detection Mechanisms for Digital Documents](#) to appear in SIGMOD'95 (PostScript Version).

Chang C.K., Garcia-Molina, H., [Evaluating the Cost of Boolean Query Mapping](#) (working paper version) Digital Libraries '97.

Chang, E., Garcia-Molina, H., [Reducing Initial Latency in a Multimedia Storage System](#) Third International Workshop of Multimedia Database Systems (PostScript Version).

Chang, C.K., Garcia-Molina, H., Paepcke, A., [Boolean Query Mapping Across Heterogeneous Information Sources](#) (working paper version) IEEE Transactions on Knowledge and Data Engineering 8(4):515-521, Aug, 1996.

Cousins, S.B., Paepcke, A., Winograd, T., Bier, E.A., Pier, K., [The](#)

[Digital Library Integrated Task Environment \(DLITE\)](#) (working paper version) Digital Libraries '97.

Cousins, S.B., Ketchpel, S., Paepcke, A., Garcia-Molina, H., Hassan, S., Röscheisen, M., [InterPay: Managing Multiple Payment Mechanisms in Digital Libraries](#) In Digital Libraries 95 (PostScript Version).

Crespo, A., Bier, E.A., [WebWriter: A Browser-Based Editor for Constructing Web Applications](#) Fifth International World Wide Web Conference, 1996. [Slide presentation in HTML](#).

Crespo, A., Chang, B.W., Bier, E.A., [Responsive Interaction for a Large Web Application: The Meteor Shower Architecture in the WebWriter II Editor](#) (working paper version) In Proceedings of the Seventh World-Wide Web Conference, 1997.

DLcore mailing list, [Digital Library sketch](#) Appeared in Comm. of the ACM Digital Libraries issue, 1995.

Gravano, L., Shivakumar, N., Garcia-Molina, H., [dSCAM: Finding Document Copies Across Multiple Databases](#) Proceedings of 4th International Conference on Parallel and Distributed Information Systems, 1996 (PostScript Version).

Gravano, L., Chang, C.K., Garcia-Molina, H., Paepcke, A., [STARTS: Stanford Proposal for Internet Meta-Searching](#) In Proceedings of the International Conference on Management of Data, 1997 (PostScript Version).

Gravano, L., Garcia-Molina, H., [Generalizing GLOSS to Vector-Space Databases and Broker Hierarchies](#) Appeared in VLDB'95, September 1995 (PostScript Version).

Koller, D., Shoham, Y., [Information agents: A new challenge for AI](#) IEEE Expert:8-10, June, 1996 (PostScript Version).

James, F., [Presenting HTML Structure in Audio: User Satisfaction with](#)

[Audio Hypertext](#), In ICAD '96 proceedings, Nov. 1996 (PostScript Version).

James, F., [Presenting HTML Structure in Audio: User Satisfaction with Audio Hypertext](#), CSLI Technical Report Number 97-201. Extended version of ICAD publication.

James, F., [AHA: Audio HTML Access](#) In WWW6 proceedings, April 1997.

Kamiya, K., Röscheisen, M., Winograd, T., [Grassroots: A System Providing A Uniform Framework for Communicating, Structuring, Sharing Information, and Organizing People](#) CHI96, WWW96.

Ketchpel, S.P., Garcia-Molina, H., Paepcke, A., [Shopping Models: A Flexible Architecture for Information Commerce](#) (working paper version) Digital Libraries '97.

Ketchpel, S., Garcia-Molina, H., Paepcke, A., Hassan, S., Cousins, S., [UPAI: A Universal Payment Application Interface](#) (working paper version) In USENIX 2nd e-commerce workshop, 1996.

Ketchpel, S., Garcia-Molina, H., [Making Trust Explicit in Distributed Commerce Transactions](#) International conference on Distributed Computing Systems, 1996 (PostScript Version).

Ketchpel, S., [Transaction Protection for Information Buyers and Sellers](#) To appear DAGS '95, Electronic Publishing and the Information Superhighway] .

Kohavi, R., Sahami, M., [Error-Based and Entropy-Based Discretization of Continuous Features](#) In Second International Conference on Knowledge Discovery in Databases, 1996 (PostScript Version).

Kohavi, R., Sahami, M., [Error-Based and Entropy-Based Discretization of Continuous Features](#) Second International Conference on Knowledge Discovery in Databases, 1996 (PostScript

Version).

Koller D., Sahami, M., [Toward Optimal Feature Selection](#) ICML-96, 1996 (PostScript Version).

Lynch, C., Garcia-Molina, H., et al, [IITA Digital Libraries Workshop report](#).

Paepcke, A., [Searching is Not Enough: What We Learned On-Site](#) D-Lib Magazine, May, 1996.

Paepcke, A., [Information Needs in Technical Work Settings and their Implications for the Design of Computer Tools](#) (working paper version) Computer Supported Cooperative Work: The Journal of Collaborative Computing, 5:63-92, 1996.

Paepcke, A., Cousins, S.B., Garcia-Molina, H., Hassan, S.W., Ketchpel, S.P., Röscheisen, M., Winograd, T., [Towards Interoperability in Digital Libraries: Overview and Selected Highlights of the Stanford Digital Library Project](#) In IEEE Computer, May 1996.

Röscheisen, M., Winograd, T., [A Communication Agreement Framework of Access/Action Control](#) In Proceedings of the 1996 IEEE Symposium on Research in Security and Privacy, Oakland (PostScript Version).

Röscheisen, M., Winograd, T., Paepcke, A., [Content Rating and Other Third-Party Value-Added Applications for the World-Wide Web](#) CNRI Journal D-Lib, August 1995.

Röscheisen, M., Mogensen, C., Winograd, T., [Interaction Design for Shared World-Wide Web Annotations](#) In CHI'95. See also ComMentor Documentation and screenshots.

Röscheisen, M., Mogensen, C., Winograd, T., [A Platform for Third-Party Value-Added Information Providers: Architecture, Protocols, and Usage Examples](#) ComMentor/PCD Brio Technical Report. Includes technical specifications. [Slides from WWW](#)

[conference talk](#), [ComMentor Documentation](#).

Röscheisen, M., Mogensen, C., Winograd, T., [Beyond Browsing: Shared Comments, SOAPs, Trails and On-line Communities](#) appeared in WWW'95, Darmstadt. [ComMentor Documentation](#) by M. Röscheisen, C. Mogensen and T. Winograd.

Sahami, M., Hearst, M., Saund, E., [Applying the Multiple Cause Mixture Model to Text Categorization](#) In Proceedings of the Thirteenth International Conference on Machine Learning, pp. 435-443. Morgan Kaufmann, 1996 (PostScript Version).

Sahami, M., [Learning Limited Dependence Bayesian Classifiers](#) Second International Conference on Knowledge Discovery in Databases, 1996 (PostScript Version).

Shivakumar, N., Garcia-Molina, H., [The SCAM Approach to Copy Detection in Digital Libraries](#) Diglib Magazine, November 1995.

Shivakumar, N., Garcia-Molina, H., [SCAM: A Copy Detection Mechanism for Digital Documents](#) Accepted to Digital Libraries '95 (PostScript Version).

Wang Baldonado, M.Q., Cousins, S.B., [Addressing heterogeneity in the networked information environment](#) (working paper version) The New Review of Information Networking, Volume 2, 1996, p. 83-102.

Wang Baldonado, M.Q., Winograd, T., [SenseMaker: An Information-Exploration Interface Supporting the Contextual Evolution of a User's Interests](#) (working paper version) In Proceedings of the Conference on Human Factors in Computing Systems, 1997.

Shivakumar, N., Garcia-Molina, H., Building a Scalable and Accurate Copy Detection Mechanism In DL '96 Proceedings, 1996.

Yan, T.W., Jacobsen, M., Garcia-Molina, H., Dayal, U., From User Access Patterns to Dynamic Hypertext Linking In Fifth World-Wide Web Conference, 1996.



Comments to [Susan Harum](#) at the University of Illinois
Last updated 5/1/97

[Outline](#)[Achievements](#)[Sponsors](#)[Links](#)[Search](#)[Home](#)

A variety of research supports the ability of the Informedia Digital Library to index and retrieve video, audio and image materials:

Publications & Reports

- [Publications](#)
- [NSF Reports](#)

Research Areas

- [Image Understanding](#)
- [Natural Language Processing](#)
- [Speech Recognition](#)
- Human Compute Interface
- Networked Data Transport
- [Network Billing & Security](#)

Upcoming Conferences & Presentations

- [AAI Spring Symposium](#)

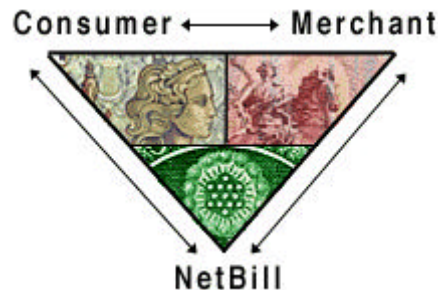
Copyright 1996 Carnegie Mellon University



The NetBill Project

- ◆ Overview
- ◆ News
- ◆ Publications
- ◆ Technical Partners
- ◆ Project Members
- ◆ Licensing Information
- ◆ Job Opportunities
- ◆ Commerce Resources

A dependable, secure, and economical payment method for purchasing digital goods and services through the Internet.



The NetBill electronic commerce project at Carnegie Mellon's [Information Networking Institute](http://www.ini.cmu.edu/) is researching design issues of highly survivable and secure distributed transaction processing systems, as well as accounting and access control for digital libraries. NetBill is addressing these issues by developing the protocols and software to support network-based payment for goods and services over the Internet.

These protocols and software have been implemented in a test system, currently in its Alpha trial, on the Carnegie Mellon campus. This system enables consumers and merchants to communicate directly with each other, using NetBill to confirm and ensure security for all transactions.

We invite you to take a look at this test system at:

<http://www.netbill.com>

Please note that the NetBill functions are restricted to the Carnegie Mellon campus. However, there is plenty of information about NetBill for you to explore.

For more information about the NetBill project, please explore this web site using the links on the left of each page.

If you require further information, please contact us at support@netbill.com



All contents copyright © 1995,1996,1997 Carnegie Mellon University.

All rights reserved.

Last revision: Fri Apr 11 12:36:17 EDT 1997



Next: [Task Planner Agent](#) **Up:** [Task Planning Agents in](#) **Previous:** [Introduction](#)

Task Planning in the UMDL Architecture

A fundamental activity in the University of Michigan Digital Library (UMDL) is connecting people/agents that need help accomplishing their tasks with people/agents that are capable of doing those tasks. A canonical example is that of a library user in search of information who needs to contact the appropriate collections which contain the information. The job of a task planning agent (TPA) is to forge these connections between agents -- to help agents team up with the right agents in a large, open, and constantly evolving network of agents.

We envision that there will be many TPAs within the UMDL. They will have in common the general role of finding resources (agents with capabilities and/or content of interest), but each will possess specific knowledge and procedures for doing so, depending both on the characteristics of the tasks that are in need of resources, and on the resources available for doing the search (e.g. monetary funds, user patience, etc.). The class of tasks that we are initially focusing on are, not surprisingly, query answering tasks. Thus, in this paper, we will consider the subclass of TPAs that are specialized for query tasks, and will focus specifically on how we have designed and built an instance of a TPA for query planning.

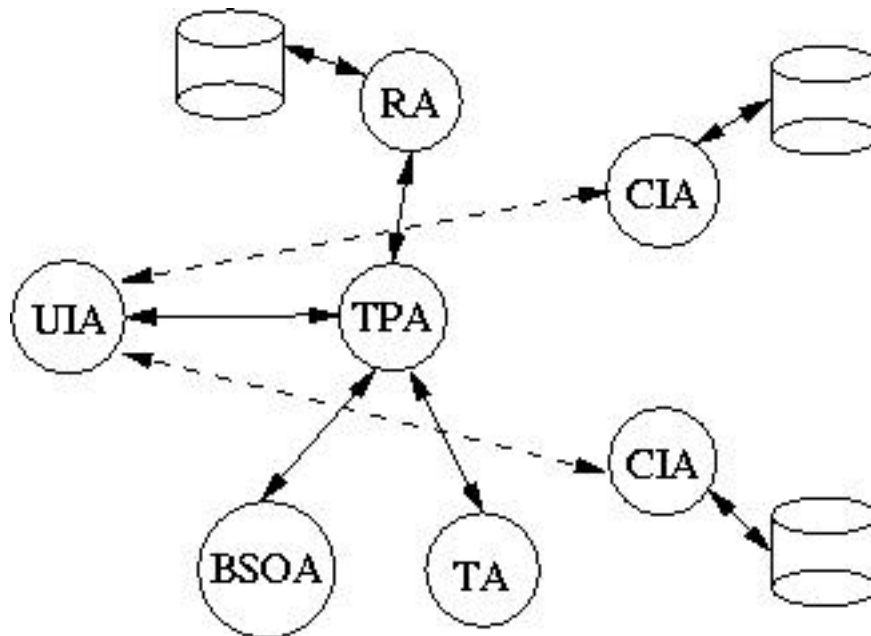


Figure 1: Schematic view of the UMDL architecture showing the agents in it and the communications between them. The solid lines represent the first part of the search, where the TPA looks for applicable collections. The dashed lines are the second part, when the TPA queries the collections that were found. Future versions will have several instantiations of each agent.

A query planning TPA is ultimately responsible for finding one or more collections based on the needs of a user; in order to do so, it communicates with many agents in the UMDL, as seen in Figure 1. The TPA receives the specification of the user's needs, along with parameters concerning task planning such as constraints on the number of collections to find or the effort to expend in finding them, from the User Interface Agent (UIA) which interacts directly with the user. The TPA needs to examine the query task to identify the characteristics of collections that might satisfy it, and from this formulate a query to the Registry Agent (RA) to see whether such collections exist. The TPA can inspect the responses from the RA and might simply forward

(a subset of) them back to the UIA -- but, more often, what is returned does not satisfy the constraints of the query task. The TPA possesses procedural knowledge about how to handle such cases, which could in turn involve enlisting the help of other agents in the network, such as agents that can broaden/narrow topics of search (Broad System of Ordering Agents (BSOAs)) and agents that can provide synonymous terms (Thesaurus Agents (TAs)). Moreover, for a candidate collection, the TPA could contact the Collection Interface Agent (CIA) associated with it, to probe more deeply into its capabilities, content, price, or availability. Ultimately, the TPA formulates a candidate set of collections for the user to contact, and returns the set to the UIA. The UIA might then proceed to contact some of these CIAs directly and ask for the particular documents, before returning the final answer to the user.

-
- Task Planner Agent
 - TPA architecture
 - TPA language
 - TPA knowledge



Next: Task Planner Agent **Up:** Task Planning Agents in **Previous:** Introduction

Jose M. Vidal
jmvidal@umich.edu
Mon Mar 11 15:41:27 EST 1996



UMDL Organization

Picture of UMDL Research Project Categories of Activities (DEA) [HERE!](#)

The above graph illustrates the interaction among several activities of our digital library project. More in-depth activities of many of these teams are represented below. Some of the small teams do not have their own web page.

The UMDL project has a continually evolving set of teams which interact to create the production system and do relevant research. Members of all of these teams meet on the first Friday of each month to report progress, discuss important issues, and get updates on events and activities of the project.

Below is a list of current teams, as well as a "master" list of active project members (does not include people who are loosely affiliated with the project). By selecting one of the teams, you will be able to see a list of members, regular meeting times, and meeting notes and other documentation.



Advanced User Interface Group

This team is primarily concerned with a state-of-the-art user centered design digital library. Their research are long term and results will be folded into the production system.



Architecture

This team develops the agent architecture and conducts research in related areas.



ColSR

ColSR, or Collection Search and Retrieval, is focused on search and retrieval functions related to the overall digital library collection.



Conspectus Definition and Registry

This team works on the development of the conspectus.



ConSR

ConSR, or Conspectus Search and Retrieval, is focused on search and retrieval functions related to the conspectus.



Intellectual Property and Economic Issues

This team is working on commerce mechanisms, intellectual property license management, agent negotiation protocols, and resource allocation issues for the UMDL.



Operating Committee

This team is comprised of team leaders from all areas.



Use and Evaluation in Education

This team is responsible for deployment and evaluation of the production system in high schools and public libraries.



User Interface Design and Evaluation

This team is primarily concerned with the design of the current production system.



Master List of all active members of the project

This alphabetical master lists provides personal web sites and affiliations for UMDL project members.



Project Partners

This lists partners, from industry to educational, who are involved in the project.

[Return to the Main Page](#)

Comments or questions may be sent to: UMDL.INFO@umich.edu



Stanford University Digital Libraries Project

The Stanford Digital Libraries project is one participant in the 4-year, \$24 million Digital Library Initiative, started in 1994 and supported by the [NSF](#), [DARPA](#), and [NASA](#). In addition to the ties with the [five other universities](#) that are part of the project, Stanford also has a large number of [industrial partners](#). Each university project has a different angle of the total project, with Stanford focusing on **interoperability**.

Our collection is primarily computing literature. However, we also have a strong focus on networked information sources, meaning that the vast array of topics found on the World Wide Web are accessible through our project as well. At the heart of the project is the [testbed running the "InfoBus" protocol](#), which provides a uniform way to access a variety of services and information sources through "proxies" acting as interpreters between the InfoBus protocol and the native protocol.

With the InfoBus protocol running under the hood, a variety of user level applications provide powerful ways to [find information](#), using cutting-edge [user interfaces](#) for direct manipulation or through [Agent technology](#). A second area of focus for the Stanford Digital Library Project is the [legal and economic issues](#) of a networked environment.

[PROJECTS](#)

See the entire list, or jump directly to projects related to [information finding](#), [user interfaces](#), [legal and economic issues](#), [the testbed](#), or [agents](#).

[DOCUMENTS](#)

A collection of introductory information, and our [publications](#), our [working papers](#), our [presentations](#), our [mailing archives](#), and our [project reports](#).

[INFO RESOURCES](#)

A collection of pointers to digital library-related resources, both at Stanford and elsewhere.

[SEMINARS](#)

A schedule of our weekly Digital Library seminar, which meets Mondays at 4:15 in Gates 104.

[SOFTWARE](#)

A collection of software developed for and used by the Stanford Digital Library project.

[PEOPLE](#)

A list of the Stanford faculty, staff, and student participants and industrial partners.




May 2, 1997

[Presentations](#) from our recent funders' site visit give an excellent overview of current research from the Stanford Digital Libraries crew. Our [publications page](#) has also been recently updated to more than 3 dozen reports!



Quick Tabs to Projects: [GLOSS](#) -- [Query Translation](#) -- [SenseMaker](#) -- [FAB](#) -- [STARTS](#) -- [Grassroots](#) -- Machine Learning -- [BackRub](#) -- [ComMentor](#) -- [R-Manage](#) -- [InterPay](#) -- Distributed Transactions --

[InterOp Protocol](#) -- Z Server -- Proxy Generator -- [Infobus Socket Interface](#) -- [JYLU](#) -- [DLITE](#) -- [Audio HTML Access](#) -- WebWriter --
[Interbib](#) -- [SCAM](#) -- [COPS](#)

<p><u>Stanford Digital</u> <u>Library</u> Testbed Development</p> <p>Department of Computer Science Stanford University Stanford, CA</p>	
---	--

What is the Stanford Digital Library Testbed?

The Stanford Digital Library testbed is our platform for experimentation with interoperation among online services. Our basic approach is to use **distributed objects** to allow integrated access to heterogeneous services across networks. The distributed approach allows the interaction of processes on different machines, with different architectures, implemented in different languages. We use **CORBA** to provide communication between remote processes. In particular, we use Xerox PARC's **ILU**, a free implementation of a CORBA superset. It offers language bindings for C++, C, CommonLisp, Python and Modula-3. We use the interpreted, object-oriented language Python for most of our development work.

For more information, see:

CORBA

- Information from the [OMG](#), including a [Manual](#)
- [Common Object Services](#) developed at Stanford

ILU

- [Xerox PARC's ILU Home Page](#)
- The current [ILU Manual](#)

- Information about the [Stanford installation](#)
- A [technical performance evaluation](#) of ILU, HTTP, and basic TCP

Python

- [The Python Language Home Page](#)
 - Information about the [Stanford installation](#)
-

What Protocol does the Testbed Use?

We have developed the **Digital Library Interoperation Protocol (DLIOP)** for information access and retrieval. It is an asynchronous protocol, providing robustness in the face of network or server outages. Moreover, it also gives the programmer a high degree of control over where and when information objects are materialized, affecting tradeoffs of space and cost vs. time. This protocol has been adopted by other participants of the Digital Library Initiative, including University of Michigan and University of California at Santa Barbara.

For more information, see:

- A PostScript version of the [Full Protocol Specification](#)
 - A PostScript version of a [presentation describing DLIOP](#)
 - [A full CORBA IDL specification of DLIOP](#), or the [specification in ILU's ISL](#)
-

How Can I Use the Stanford Testbed?

Even if you're not local to Stanford, there are two simple ways of accessing the InfoBus from a remote site. Both use the DLIOP protocol. **IBClient**, the first method, accesses the InfoBus through ILU calls and is therefore a full-functionality client. It requires client sites to have ILU or another CORBA implementation installed. The second

alternative, the [InfoBus Socket](#) delivers the DLIOP calls via an ASCII stream over a socket. It does not require a CORBA implementation, but it is of limited functionality.

We include [code for the IBClient](#) and [code for the InfoBus Socket](#). The IBClient example is written in Python and thus requires your machine to have [Python](#) installed. If you do not have Python, you can use the example to build your own client in the language of your choice. The example represents a minimum, bare-bones client-side implementation. See the DLIOP documentation for additional facilities that can be added to clients. As an example, first download all of the files from the IBClient directory (the subdirectory CVS is not required). Then, try typing *ibclient.py WebCrawler 'digital library'* to see the titles of WebCrawler searches for those keywords. The InfoBus Socket is written in C++. It mimics the DLIOP calls in syntax, but it delivers them through UNIX sockets. See the [description of how it works](#).

We also have CORBA interface specification files for the DLIOP protocol.

- [IDLInterchange.idl](#)
- [IDLInterchange.isl](#)

Information of Interest to local Stanford Developers

- [ILU -- Our installation and examples.](#)
 - [Python -- Our installation of Python programming language](#)
 - [CVS -- Our use of CVS in the testbed.](#)
 - [Various Manuals \(CVS, Python, ILU\)](#)
-



Information Finding Projects in the Stanford Digital Library

One of the major research thrusts of the Stanford Digital Library project is helping users to find information. We have initiated a number of projects in this area, most related to our over-arching theme of interoperability. We have looked at ways that search tools can be used across multiple sources that use different syntaxes or languages. We have also looked at tools to provide statistical or collaborative filtering to locate relevant articles.

FAB

FAB is an adaptive multi-agent information retrieval system which finds interesting pages on the web.

"An Adaptive Agent for Automated Web Browsing"

- Marko Balabanovic

GLOSS

The Glossary Server of Servers (GLOSS) project is designed to locate relevant information sources for your query.

"Generalizing GLOSS to Vector-Space Databases and Broker Hierarchies"

- Luis Gravano
-

Query Translator

Databases have different query syntax and different capabilities, even for simple Boolean queries. Translation allows a single query to be mapped into the native format appropriate for each database.

- [Chen-Chuan K. Chang](#)
-

SenseMaker

SenseMaker helps users iteratively reformulate their information needs through multi-dimensional organizing and active gathering of search results.

" [SenseMaker: An Information-Exploration Interface Supporting the Contextual Evolution of a User's Interests](#)"

- [Michelle Q Wang Baldonado](#)
-

Grassroots

Groupware for information finding, combines mail, news, and web in a single environment with distribution lists

" [Grassroots: A System Providing a Uniform Framework for Communicating, Structuring, Sharing Information, and Organizing People](#)"

- [Kenichi Kamiya](#)
 - [Martin Röscheisen](#)
-

STARTS: Stanford Protocol Proposal for Internet Retrieval and Search

A set of informal standards negotiated among the major search vendors and users to facilitate interoperation.

- [Chen-Chuan K. Chang](#)
- [Hector Garcia-Molina](#)

- [Luis Gravano](#)
 - [Andreas Paepcke](#)
-

Machine Learning for Information Retrieval

Statistical AI techniques allow the extraction of minimal sets of meaningful search terms

["Toward Optimal Feature Selection"](#)

- [Mehran Sahami](#)
 - [Daphne Koller](#)
-

[BackRub](#)

BackRub is a web crawler which is designed to store the connection graph for the web. In other words BackRub stores which pages every web page links to. Currently we are developing techniques using this link data to improve web search engines as well as understand the structure of the web.

- [Larry Page](#)
-

[ComMentor](#)

Third-Party Annotations on web pages provide for ways to share information, rate content, and keep notes

["A Platform for Third-Party Value-Added Information Providers: Architecture, Protocols, and Usage Examples"](#)

- [Martin Röscheisen](#)
 - [Christian Mogensen](#)
 - [Terry Winograd](#)
-

[InterOp Protocol](#)

The heart of the "InfoBus", this protocol describes access methods to search collections, acquire results, and find out about

Sources.

- [Steve Cousins](#)
- [Prof. Hector Garcia-Molina](#)
- [Scott Hassan](#)
- [Andreas Paepcke](#)

[SCAM: The Stanford Copy Analysis Mechanism](#)

Making a perfect digital copy of a copyrighted work is easy in a networked world. How can the intellectual property rightsholders be protected? By detecting attempted distribution of illegal copies. Duplicate detection has other uses in information finding as well. An earlier, related project was known as COPS: The Copyright Protection Scheme.

["Building a Scalable and Accurate Copy Detection Mechanism"](#)

- [Prof. Hector Garcia-Molina](#)
- [Narayanan Shivakumar](#)

[InterBib](#)

InterBib is a tool for maintaining bibliographic information. Capable of reading from and writing to many different formats, it acts as a unified, searchable repository of bibliographic records.

[Information on InterBib](#)

- [Andreas Paepcke](#)





User Interface Projects in the Stanford Digital Library

Too often the power of a search engine goes untested because users don't know how to exploit the advanced (or even basic) features. The use of a browser front-end has eased platform independent rapid prototyping, allowing a wide variety of services such as information clustering, annotating, and re-distributing via the WWW. One project even uses a web application to help create web applications! But the web does have drawbacks, such as being largely inaccessible to blind users (hear our audio interface!) and limiting the types of possible interaction. Therefore, our DLITE interface uses a direct manipulation metaphor of iconic representations, rather than relying on CGI forms.

[SenseMaker](#)

SenseMaker helps users iteratively reformulate their information needs through multi-dimensional organizing and active gathering of search results.

" [SenseMaker: An Information-Exploration Interface Supporting the Contextual Evolution of a User's Interests](#)"

- [Michelle Q Wang Baldonado](#)
-

[DLITE: A Digital Library Interface](#)

A direct manipulation user interface designed to support user tasks, to smoothly integrate the results of many services, to handle services of widely-varying time scales, to be extensible, and to support sharing and reuse.

["The Digital Library Integrated Task Environment \(DLITE\)"](#)

- [Steve Cousins](#)
-

[Grassroots](#)

Groupware for information finding, combines mail, news, and web in a single environment with distribution lists

["Grassroots: A System Providing a Uniform Framework for Communicating, Structuring, Sharing Information, and Organizing People"](#)

- [Kenichi Kamiya](#)
 - [Martin Röscheisen](#)
-

[ComMentor](#)

Third-Party Annotations on web pages provide for ways to share information, rate content, and keep notes

["A Platform for Third-Party Value-Added Information Providers: Architecture, Protocols, and Usage Examples"](#)

- [Martin Röscheisen](#)
 - [Christian Mogensen](#)
 - [Terry Winograd](#)
-

Audio Interfaces to HyperText

The structure of a document is captured in HTML/SGML tags which most browsers map to visual display characteristics. We are seeking ways in which this structural information can be conveyed in audio format for blind users or users connecting via telephone.

[CSLI Annual Report](#)

- [Frankie James](#)
 - [Prof. Terry Winograd](#)
-

WebWriter

WebWriter is a direct manipulation Web page editor that allows users to create new web pages, including advanced features such as tables, without knowing HTML or CGI.

["WebWriter: A Browser-Based Editor for Constructing Web Applications"](#)

- [Arturo Crespo](#)
-

[RManage/FIRM](#)

Interoperable rights management is one of the service layers that the current Internet is still lacking. FIRM defines a platform for "smart contracts" that is based on a computational reification of contract law; it is realized as part of a novel, network-centric architecture for managing control information that generalizes previous models centered around clients or servers.

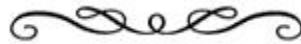
["A Network-Centric Design for Relationship-based Rights Management"](#)

- [Martin Röscheisen](#)
 - [Prof. Terry Winograd](#)
-





Digital Library Integrated Task Environment



The Digital Library Integrated Task Environment (DLITE) is an experimental, direct-manipulation interface to information objects and services. Information services are accessed via the InfoBus, and are presented to the user as components in workcenters.

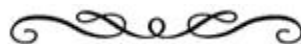


For More Information...

- ▶ [Interface Details](#)
- ▶ [Interface Architecture paper](#)
- ▶ [Summary of interface goals \(CHI '96 paper\)](#)
- ▶ [List of DLITE Components](#)
- ▶ [Screen Shots](#)



DLITE is implemented as a distributed, client/server application. The server is written in Python, and clients have been written in Python/Tk and Java/AWT. DLITE makes extensive use of the Stanford InfoBus for search and query translation. We have completed a pilot study of the interface, and are continuing to test various aspects of it as well.



Credits...

DLITE is the PhD project of Steve Cousins. The following people have helped to build or design various aspects of the system:

- Scott Hassan
- Alan Steremberg
- Terry Winograd
- Ken Pier
- Eric Bier
- Andreas Paepcke
- Mark Mortensen

DLITE Components

Basic Elements

 document

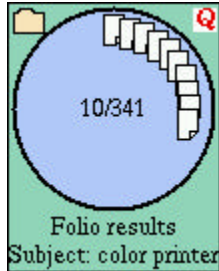
Parent: component.component

In File: documents.py

 query

Parent: component.component

In File: queries.py



collection

Parent: component.component

In File: collection.py



service

Parent: component.component

In File: services.py



epersComponent

Parent: component.component

In File: epersComponent.py

Constructors

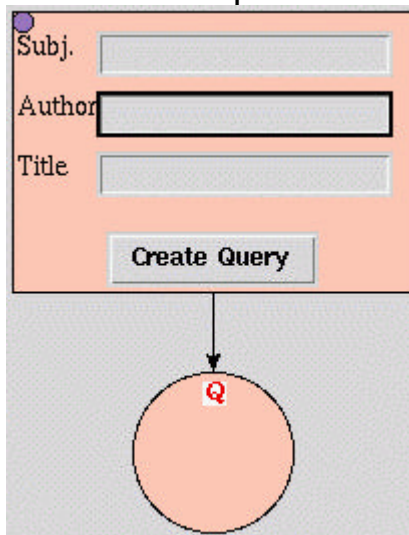
These components are used to create new components of various types.

CollectionConstructor

Parent: component.component

In File: collection.py

Creates: collection.collection

DocumentConstructor**Parent:** component.component**In File:** documents.py**Creates:** documents.document**LabelConstructor****Parent:** component.component**In File:** labelComponent.py**Creates:** labelComponent.label**LabelConstructor****Parent:** component.component**In File:** labelComponent.py**Creates:** labelComponent.label**LabelMaker****Parent:** component.component**In File:** labelComponent.py**Creates:** labelComponent.label**Note:** Experimental variant of LabelConstructor**QueryConstructor****Parent:** component.component**In File:** expertqueries.py**Creates:** queries.query**QueryConstructor****Parent:** component.component**In File:** queries.py**Creates:** queries.query

ShCollectionConstructor

Parent: CollectionConstructor

In File: collection.py

Creates: collection.shCollection

SourceConstructor

Parent: component.component

In File: sources.py

Creates: sources.source

TaskConstructor

Parent: component.component

In File: taskObject.py

Creates: taskObject.taskObject

documentConstructor

Parent: component.component

In File: netscapeSVC.py

Creates: documents.document

Default Components

Inbox

Parent: component.component

In File: inboxComponent.py

registry

Parent: component.component

In File: registry.py



trashCan

Parent: component.component

In File: trashComponent.py

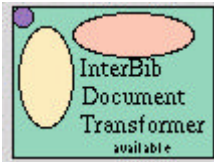
Services



source

Parent: services.service

In File: sources.py



InterBibDocTransformer

Parent: services.service

In File: bibtransSVC.py



SenseMaker

Parent: services.service

In File: sensemaker.py



bibSVC

Parent: services.service

In File: bibSVC.py



fab

Parent: component.component

In File: fab.py

multiSearch

Parent: component.component

In File: multiSearch.py



netscape_uploader

Parent: component.component

In File: upload.py

summarizer

Parent: services.service

In File: summarizer.py



oracleSummarizer

Parent: summarizer

In File: summarizer.py



parcSummarizer
Parent: summarizer
In File: summarizer.py

translator

Parent: services.service
In File: translator.py



prescriptTranslator
Parent: translator
In File: translator.py

publisher

Parent: component.component
In File: pubSVC.py



scamSVC
Parent: component.component
In File: scamService.py

sdiSVC

Parent: component.component
In File: sdi.py

Collections

glossResultSet

Parent: resultSet.resultSet
In File: glossResultSet.py

resultSet

Parent: collection.collection
In File: resultSet.py

sdiResultSet

Parent: resultSet.resultSet
In File: sdi.py

sensemakerCollection

Parent: collection.collection

In File: sensemakerCollection.py
shCollection
Parent: collection
In File: collection.py
spiralCollection
Parent: collection
In File: collection.py
triggeredShCollection
Parent: shCollection
In File: collection.py

Miscellaneous

label
Parent: component.component
In File: labelComponent.py
taskObject
Parent: component.component
In File: taskObject.py
tt
Parent: component.component
In File: testTable.py

Holophrasting Components

dclInstructions
Parent: holoframe.verticalHoloframe
In File: holophrastic.py
instWithComponent
Parent: instructionStep
In File: holophrastic.py
instructionStep
Parent: component.component
In File: holophrastic.py
verticalHoloframe
Parent: component.component

In File: holoframe.py

RManage components

Bank

Parent: sources.source

In File: bankComp.py

certificate

Parent: component.component

In File: certificates.py

certifier

Parent: component.component

In File: certifiers.py

check

Parent: documents.document

In File: checkConstructor.py

cpctComponent

Parent: component.component

In File: cpctComponent.py

CommpactFinder

Parent: component.component

In File: cpctComponent.py

PersonFinder

Parent: component.component

In File: epersComponent.py

formsSource

Parent: sources.source

In File: formsSource.py

loginBox

Parent: component.component

In File: login.py

notifier

Parent: collection.shCollection

In File: notifier.py

offerCreator

Parent: services.service

In File: offerCreator.py
promiseComp
Parent: component.component
In File: promComponent.py
rightComp
Parent: promiseComp
In File: promComponent.py
obligationComp
Parent: promiseComp
In File: promComponent.py
PromiseFinder
Parent: component.component
In File: promComponent.py
relsBoxComp
Parent: component.component
In File: relsBox.py

RManage constructors

CertificateConstructor
Parent: component.component
In File: certificates.py
MessageConstructor
Parent: component.component
In File: certificates.py
CertifierConstructor
Parent: component.component
In File: certifiers.py
CheckConstructor
Parent: component.component
In File: checkConstructor.py
CompactConstructor
Parent: CompactFinder
In File: cpctComponent.py
PersonConstructor
Parent: PersonFinder

In File: epersComponent.py
PromiseConstructor

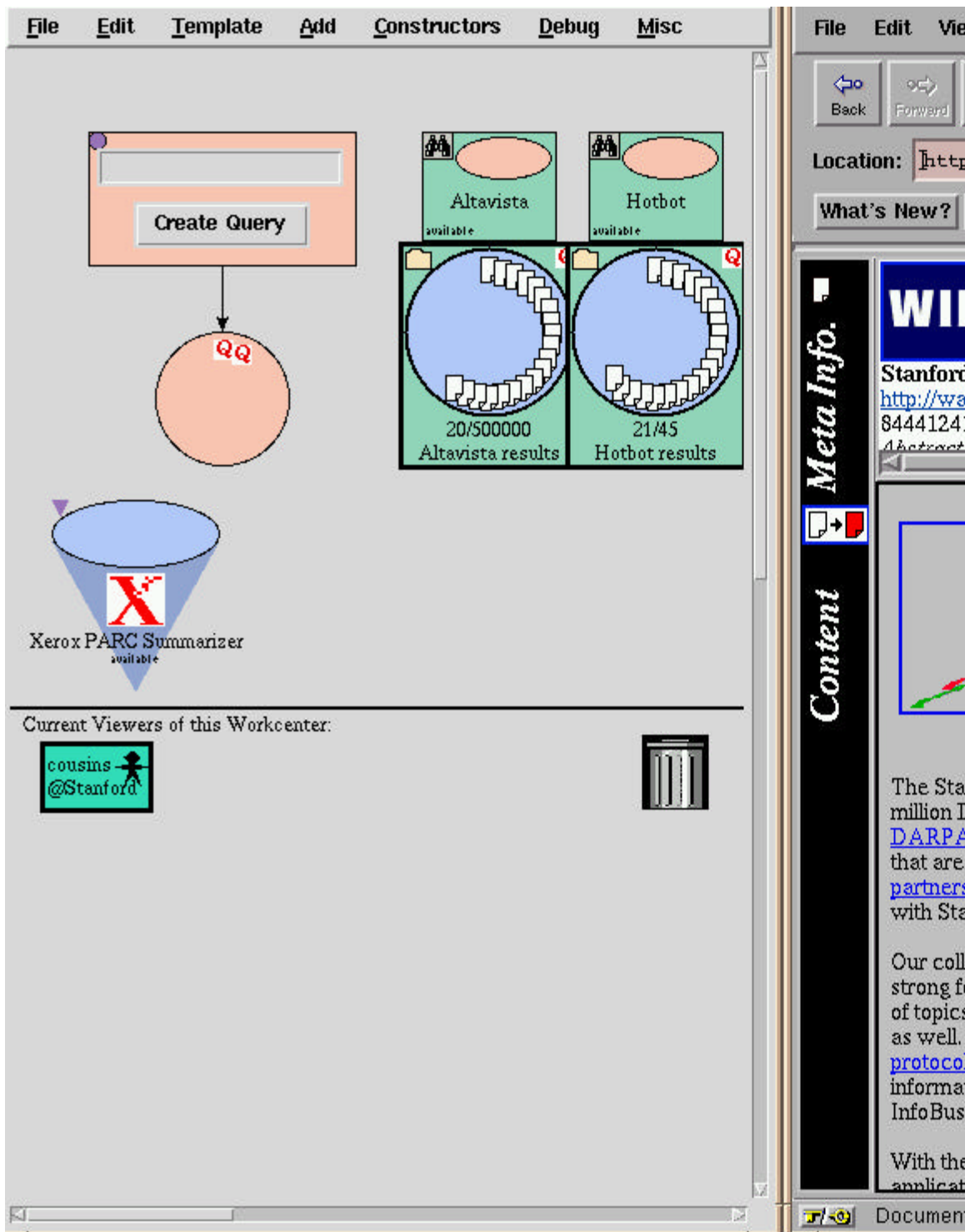
Parent: PromiseFinder

In File: promComponent.py

DLITE Screen Shots

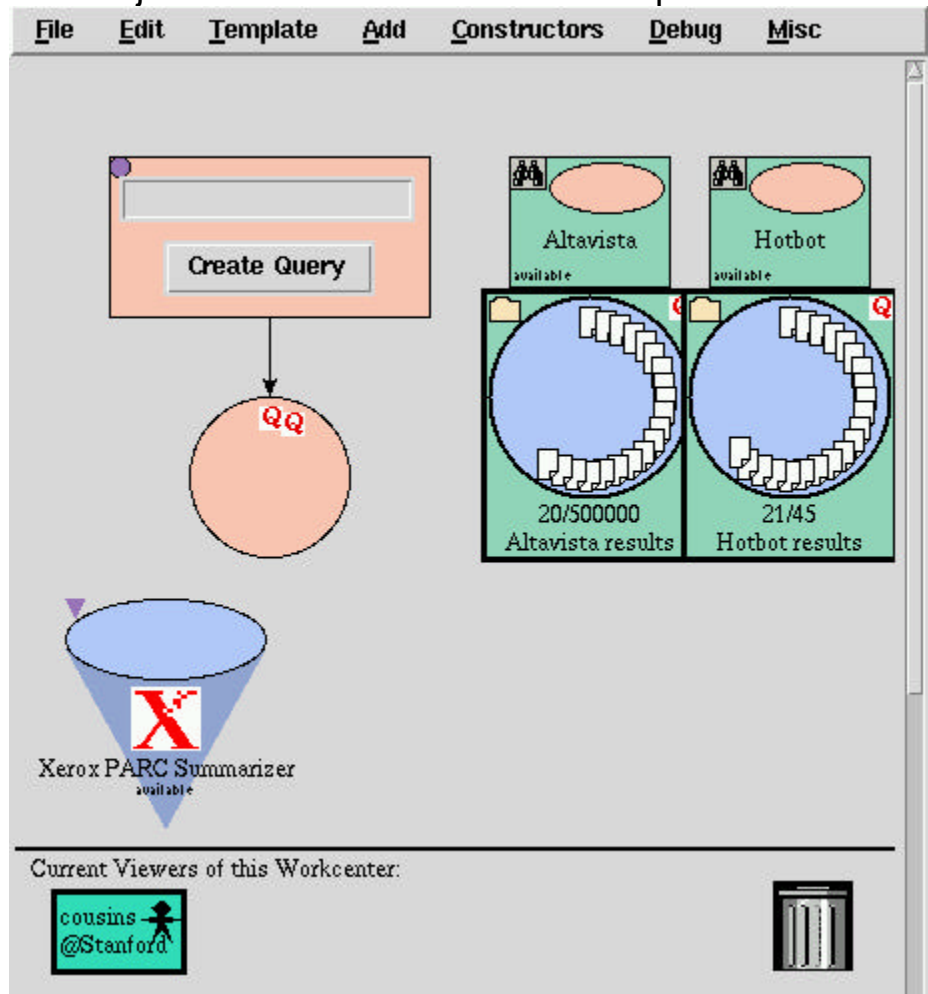
The whole screen

This image shows DLITE running next to a Netscape browser.



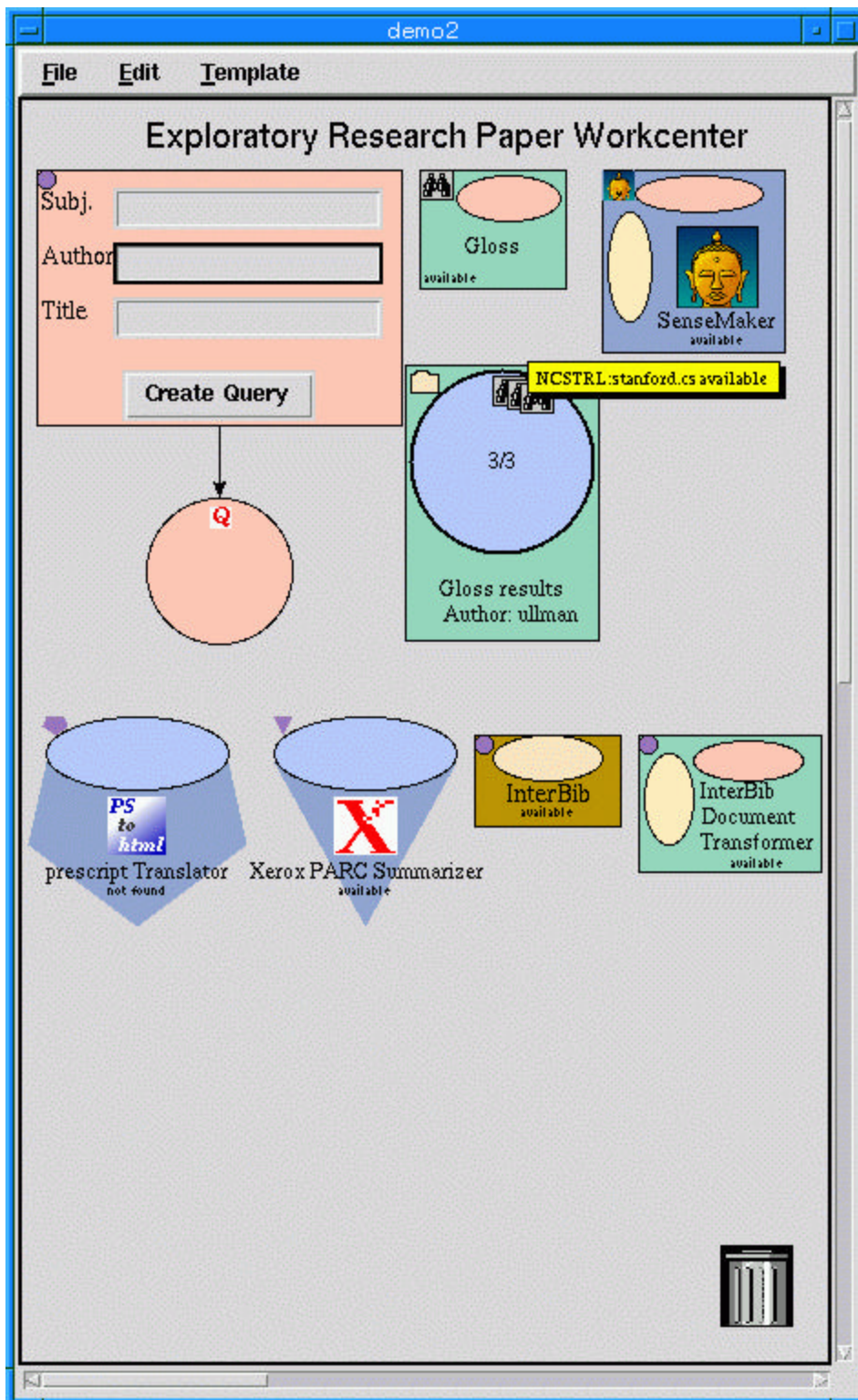
Simple search

Here is just the DLITE window for a simple search task.

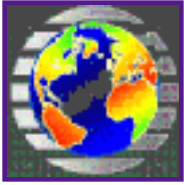


Exploratory Research Workcenter

Here is a workcenter for doing exploratory research. Services are present for source-finding, sense-making, postscript-to-ascii translation, text summarization, and bibliography generation.

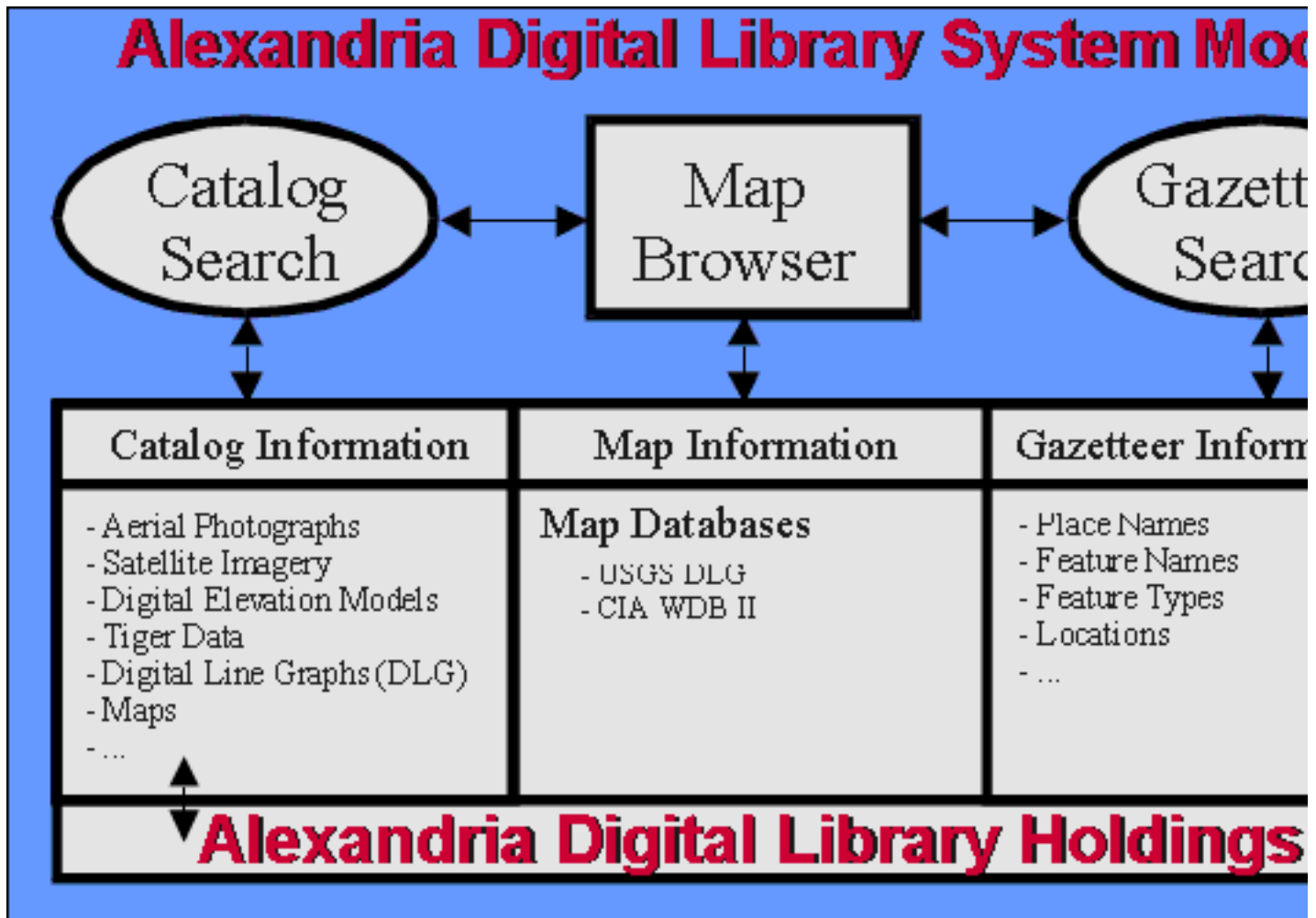


[Steve Cousins](#)



Tutorial Table of Contents

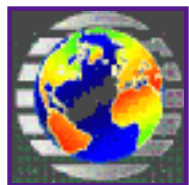
◀ Prev Next ▶



Conceptual model of the Alexandria Web interface

- Conventions
- Session / System Setup
- Map
- Gazetteer
- Catalog
- Overview of Current Holdings
- Walkthroughs (Example Sessions)
- Feedback
- Technical Reference
- Acknowledgements

1996 ANNUAL REPORT Table of Contents



Alexandria Digital Library

1 PROJECT SUMMARY

2 SIGNIFICANT EVENT

3 OVERVIEW OF PROGRESS IN LIBRARY DEVELOPMENT

4 CURRENT STATUS OF TESTBED DEVELOPMENT

4.1 Architecture of the Testbed System

4.2 Catalog The Component

4.2.1 Basic Metadata: USMARC and FGDC Standards

4.2.2 The ADL Gazetteer

4.2.3 Other Catalog Issues

4.3 Collections

4.4 The User Interface

4.4.1 User Interface Issues

4.4.2 User Interface Implementation

4.5 Image Processing

4.5.1 Browsing and Progressive Delivery

4.5.2 Texture-Based Retrieval

4.6 Parallel Processing

4.7 Computing Support for the Testbed: Hardware and Communications

4.7.1 Current Computing Equipment

4.7.2 Current Networking Support

4.7.3 Current Storage Support

4.7.4 Current "Other" Hardware and Software Support

4.7.5 Equipment and Facilities Needs

5 RESEARCH ACTIVITIES AND PROGRESS

5.1 Library Team

5.1.1 Metadata and Catalog Interoperability

5.1.2 Alexandria Atlas Subteam

5.2 Interface Design and Evaluation Team

5.3 Information Systems Team

5.4 Image Processing Team

5.5 Performance and Parallel Processing Team

6 COMPARISON OF ACTUAL AND PREVIOUSLY PLANNED R&D ACTIVITIES

6.1 Previously Planned R&D Activities

6.2 Actual Versus Agreed R&D Activities

7 PLANNING FOR AN OPERATIONAL DIGITAL LIBRARY

8 MANAGEMENT REPORT

8.1 Organizational Structure

- 8.2 Interactions and Levels of Activity
- 8.3 Personnel
 - 8.3.1 Professional Staff
 - 8.3.2 Faculty Investigators
 - 8.3.3 Graduate Students Associated with the Project

9 EXTERNAL RELATIONS AND INTERACTIONS

- 9.1 Advisory Board
- 9.2 Third DLI meeting at UCSB
- 9.3 Alexandria Metadata Workshop
- 9.4 First Alexandria Design Review
- 9.5 Interactions with Old and New Partners
 - 9.5.1 California Environmental Resources Evaluation System (CERES)
 - 9.5.2 Central Imagery Office (CIO)/National Imagery and Mapping Agency(NIMA)
 - 9.5.3 Central Intelligence Agency (CIA)
 - 9.5.4 CIESIN
 - 9.5.5 Excalibur/ConQuest
 - 9.5.6 Defense Mapping Agency
 - 9.5.7 Digital Equipment Corporation (DEC)
 - 9.5.8 ERDAS
 - 9.5.9 Environmental Systems Research Institute (ESRI)
 - 9.5.10 Hughes
 - 9.5.11 Library of Congress
 - 9.5.12 Lockheed/Space Imaging
 - 9.5.13 NASA
 - 9.5.14 Oracle
 - 9.5.15 O2
 - 9.5.16 San Diego Supercomputer Center (SDSC)
 - 9.5.17 Sierra Nevada Ecosystem Project (SNEP)
 - 9.5.18 SPOT Image
 - 9.5.19 The Analytic Science Corporation (TASK)
 - 9.5.20 United States Geological Survey (USGS)
 - 9.5.21 United States Navy, Stennis
 - 9.5.22 United States Navy, San Diego (NRAD)
 - 9.5.23 Earth Data Analysis Center (EDA), University of New Mexico
 - 9.5.24 Utah State University/Mojave Database Cooperative
 - 9.5.24 Xerox
- 9.6 Interoperability Agreements and Activities
 - 9.6.1 Stanford DLI Project
 - 9.6.2 Berkeley DLI Project
 - 9.6.3 Illinois DLI Project
 - 9.6.4 CMU DLI Project
- 9.7 Digital Libraries for School Students
- 9.8 Visits and Demonstrations
- 9.9 Talks and lectures
- 9.10 Professional interactions with DL Community
- 9.11 Publicity

10 EDUCATIONAL ACTIVITIES

11 PUBLICATIONS

- 11.1 Abstracts of Selected Articles
 - 11.1.1 Testbed Research and Development
 - 11.1.2 Information Systems Research
 - 11.1.3 Image Processing Research
 - 11.1.4 Parallel Processing Research

Building the Interspace: Digital Library Infrastructure for a University Engineering Community

Bruce Schatz, Principal Investigator
University of Illinois DLI project
dli@uiuc.edu

DLI Project-Wide Workshop
November 9, 1995 Santa Barbara, CA

Research on the Net

- The Past: Access

The Net fetches documents

- The Present: Organization

The Net searches repositories

- The Future: Analysis

The Net correlates information

- From the Internet (data transmission)
- to the Interspace (information manipulation)

Project Goals

- Semantic Federation (research)
- Distributed Repositories (infrastructure)
- Scientific Literature (testbed)
- evaluate large testbed
- perform technology research

Organizations

- Testbed

Grainger Engineering Library Information Center

(part of University Library, *UL*)

- Infrastructure

NCSA Software Development Group

(National Center for Supercomputing Applications)

- Evaluation

Graduate School of Library & Information Science (*GSLIS*) + Sociology, Economics

- Technology

GSLIS + Computer Science (*CS*), NCSA

Principal Investigators

- Bill Mischo, UL, Testbed Lead
- Tim Cole, UL, testbed collection
- Joseph Hardin, NCSA, Infrastructure Lead
- Beth Frank, NCSA, infrastructure software
- Ann Bishop, GSLIS, Testbed Evaluation
- Bruce Schatz, GSLIS, Technology Research

- executive committee: UL, NCSA, GSLIS, CS

Primary Partners

- journal/magazine publishers:

IEEE Computer Society

Institute of Electrical and Electronics Engineers (IEEE)

American Society of Civil Engineers (ASCE)

American Society of Agricultural Engineers (ASAE)

American Physical Society (APS)

American Institute of Physics (AIP)

American Astronomical Association (AAS)

John Wiley & Sons

INSPEC, Compendex (Engineering Index)

- testbed: SoftQuad, EBT, OpenText, Hewlett-Packard
- infrastructure: OCLC, CNRI, Spyglass, Microsoft

DLI Collaborators

- Stanford interoperability experiment
- Michigan search interface, OpenText SGML
- CNRI/Cornell secure object store
- Carnegie-Mellon NetBill charging trial

???

- UC Santa Barbara GIS correlations
- UC Berkeley image processing
- Carnegie-Mellon network video

Illinois Project Groups

- Testbed

process, index, search, display SGML collection

- Infrastructure

multiview interface to Web distributed repositories

- Evaluation

usage and users of testbed with information context

- Technology

semantic retrieval, both manual & automatic

analysis environments with objects & semantics

Testbed Goals

Large Organized Collection

- SGML pipeline direct from the publishers (deposit)
- complete articles fully tagged and indexed (search)

Large Number of Users

- faculty/students around UIUC then Big Ten
- Internet interface with multiple views (display)

Careful Sociological Evaluation

- needs assessment, usability studies
- surveys, instrumentation

Testbed Status

Year 2 (Sep 95 ->)

- production pipeline for a few journals
- testing production database and components
 - Grainger Library (public terminals, usability studies)
 - Beckman Institute (physics), Computer Science Dept
- implement full Web version for deployment

Year 3: University of Illinois

Year 4: CIC Universities (midwest Big Ten)

goal is 100,000 documents & 100,000 users

Testbed Collection

- full SGML from Jan 1995 forward
- production on AIP (2000), APS, ASCE
- multi-year archive from IEEE, IEEE CS
- ISO 12083 DTD with figures, tables, equations
- publisher hands-on workshop Nov 16-17
- plans for publisher-maintained repositories
- problems with rendering scientific literature

Testbed Components

- Gathering: SGML from publishers
- Processing: normalize (federate) tags
- Indexing: store term/tag lists

- Control: VisBasic search interface
- Searching: OpenText fulltext engine
- Displaying: Panorama SGML viewer
- Fetching: Mosaic gets SGML/DTD files

federated SGML repository across the Net

Production Web Version

Client Interface

- multiple view interface (Java)
- session control across repositories

Network Gateway

- repository protocols (SQL, Z39.50)
- maintains state of search history

Server Search

- deposit canonicalized SGML documents
- index using DTDs for full-text retrieval

Multiple Views

- Different Levels of Search Interface
- Drag-and-Drop between views

- Integrates A&I (Abstracting and Indexing)
- Term Suggestion followed by Text Search
- Subject Thesaurus for coarse-grain suggest
- Concept Space for fine-grain suggest
- Visual Basic prototype, Java for multiplatform

New Web Servers

- HTTP servers evolve into Object Repositories
- NCSA Web server 2.0 released December
- Modular Steps towards Repositories
- Multiple Proctols (HTTP, Keep-Open)
- Security, Metadata Checking, Link Maintain
- Stateful Gateways support Distributed Sites
- Towards Sessions in Later Versions

Sociological Evaluation

different methodologies in granularity & scale

Needs Assessment

- ethnographic observations in libraries and labs
- focus groups and user interviews

Testbed Evaluation

- conceptual framework for evaluation
- planning for usability tests of pre-productions
- development of system instrumentation

Community for Social Studies of DLs

- Allerton Institute Oct 1995 on concepts & methods
- sessions at DL95 and DL96 on user research

Technology Research

towards the Interspace: Net correlation

- Scalable Semantic Retrieval

concept spaces and vocabulary switching

- Distributed Object Stores

secure object infrastructure (w/CNRI, Cornell)

- Analysis Environment Systems

correlation of information across repositories

Semantic Retrieval

automatic indexing of concepts

- find context of phrases within documents
- generates a concept space based on term frequency

useful for interactive searching

- given a term, can suggest other terms
- merging concept spaces supports vocabulary switching

concepts require supercomputing

- concept space for INSPEC took 1 day on SGI Challenge
- co-occurrence matrix for 400K abstracts

Analysis Environment

- objects fine-grain manipulation
- navigation & grouping path recording
- retrieval & classification concept spaces
- correlations path matching via concept spaces
- prototype in Smalltalk, ObjectStore, ILU
- application in personal info, DLI, GIS

The 21st Century: Analysis

- Beyond Search to Analysis
- Cross-Correlating Information from many sources across the Net
- The Net solves problems
- Every community has its own special library
- Every community & every person does A&I !!

Go back to the Home Page

Semantic Federation from Distributed Repositories of Scientific Literature

Bruce Schatz, Principal Investigator
University of Illinois DLI project
dli@uiuc.edu

DLI Project-Wide Workshop
November 10, 1995 Santa Barbara, CA

Levels of Federation

- Syntactic
 - connection protocols (translation gateways)
- Structural
 - field names (query normalization)
 - field values (tag normalization)
- Semantic
 - context (term co-occurrence)
 - meaning (content parsing)

Testbed Federation

- Index with Document Structure

Tag normalization for field values

- Deposit with common tags after transform

problems with sections and with authors

- Search across multiple repositories

Query normalization for field names

- Gateway maps multiple protocols

problems with distribution and definition

- Display integrates multiple views

multiple sources at multiple levels

Semantic Retrieval

- automatic indexing of concepts
 - find context of phrases within documents
 - generates a concept space based on term frequency

- useful for interactive searching
 - given a term, can suggest other terms
 - merging concept spaces supports vocabulary switching
- concepts require supercomputing
 - concept space for INSPEC took 1 day on SGI Challenge
 - co-occurrence matrix for 400K abstracts

Publishing Cycle

USER: request

LIBRARY: reference

INDEXER: classify

PUBLISHER: quality

AUTHOR: generate

- users are authors, computers are publishers
- every community has a repository
- a billion repositories on the Net !!

Vocabulary Switching

- fine-grained concept spaces

for every community and subcommunity

- user and collection modeling

choose domains for user and for search

- interactive vocabulary switching

intersect at common terms to suggest across domains

- supercomputers as time machines

personal computers same computations in 5-10 years

Switching Experiments

small-scale in molecular biology (JASIS)

- worms and flies
- 5000 documents generate each space
- 10 hours per space on a workstation
- “sperm” as connection term

large-scale in engineering (in progress)

- 3M abstracts from Compendex and Inspec
- 15 large domains of engineering (200K per space)
- 10 hours per space on a supercomputer
- “fluid dynamics” as connection term

Community Repositories

- User-driven Community Searching
 - choose topics (repositories) you know
 - choose topics you want to know
- vocabulary switch across domains
 - community specific term suggestion
- Interspace simulation with 1000 communities
 - Compendex partitioned by class codes
 - 3M abstracts and 1K spaces on Convex Exemplar

first crack in large-scale “semantic” retrieval

Computer-Assisted Indexing

- domain experts but classification amateurs
 - large community A& I is too general and too old
 - small community A& I is not consistent and much labor
- useful for interactive subject classification
 - automatic suggestions for potential classifications
 - domain expert culls list from “controlled” vocabulary
- semi-automatic support via concept spaces
 - concept dictionary of tag words from co-occurrence
 - tag frequency in documents determines classification

Building the Interspace

- every machine has its own information space
- every machine has its own concept space
- spaces for every user and every community
- search is matching selected objects
- relies on computer-assisted A& I
- analysis is merging community spaces
- vocabulary switch through graph intersect

The 21st Century: Analysis

- Beyond Search to Analysis
- Cross-Correlating Information from many sources across the Net
- The Net solves problems
- Every community has its own special library



Electronic Thesis and Dissertation Initiative

Access to the material on this WWW site
presumes you have Netscape Version 3
 and Adobe Acrobat Version 3  .

NEW! [Listservs for the ETD Initiative](#)

 [Digital Library of Electronic Theses
and Dissertations](#)

(for Scholars to search and browse)

If you are looking for on-line, electronic,
Theses and Dissertations this is the
place to go. From here you can browse
the library collection of ETDs, view the
recent additions, or search the ETDs.

Workshop and Submission Information

(for Students and Directing Faculty)

If you are looking for information regarding any part of the ETD submission process, look here. This page contains useful information that can help guide you through the entire ETD development and submission processes.

Project Description and How to Join

(for Universities and other Interested Parties)

If you would like to see how the project is progressing you might be interested in looking here.

[\[Sponsorship\]](#) [\[Table of Contents\]](#) [\[Search\]](#)

Virginia Tech <ETD> Submission Form



This form will not work properly unless you are using [Netscape Navigator 2.0](#) or better.

**If you encounter a "Broken Pipe" or a "500 Server Error":
remove any HTTP proxies or try sending only one file**

Instructions:

Please fill out the form completely. Cut and paste, from your document and into the form, as necessary. Read the [help file](#) for help on cutting and pasting your abstract and for selecting keywords. Once you are done filling out the form read the [copyright statement](#) at the bottom of the page and if you agree to it click "Preview".

Document Type:

Please select the type of document you are submitting.

- ☒ Master's Thesis
- ☐ PhD Dissertation
- ☐ Special Report
- ☐ Major Paper
- ☐ Technical Report

Name:

Please enter your name just as it appears on the title page.

Email:

Please enter your email address.

Title:

Please enter the title just as it appears on the title page.

	▲
	▼
◀	▶

Degree:

Please enter the degree you will be awarded, as in
Master of Arts, Master of Science, Doctor of Philosophy.

Department:

Please enter the name of your department.

Committee Chair:

Please enter the name of your committee chair.

Chair's email:**Committee Members:**

Please enter the names (and optionally, positions) of the remainder of
your committee.

Virginia Tech Graduate School

Electronic Submission Approval Form

Student Name:

ID#:

Degree: ☐ Bachelor's ☐ Master's ☐ Doctoral degree

Document Type: ☐ Project Report ☐ Thesis ☐ Dissertation

Document Title:

Student Agreement:

I hereby certify that, if appropriate, I have obtained and attached hereto a written permission statement from the owner(s) of each third party copyrighted matter to be included in my thesis, dissertation, or project report, allowing distribution as specified below.

I hereby grant to Virginia Tech and its agents the non-exclusive license to archive and make accessible, under the conditions specified below, my thesis, dissertation, or project report in whole or in part in all forms of media, now or hereafter known. I retain all other ownership rights to the copyright of the thesis, dissertation, or project report. I also retain the right to use in future works (such as articles or books) all or part of this thesis, dissertation, or project report.

Student and Committee Agreement:

Part A. We agree that the above mentioned document be placed in the ETD archive with the following status: *(choose one of 1, 2, 3, or 4)*

☐ 1. Release the entire work immediately for access worldwide.

___ 2. Release the entire work for Virginia Tech access only.

___ 3. Secure the entire work for patent and/or proprietary purposes for a period of one year. During this period the copyright owner also agrees not to exercise her/his ownership rights, including public use in future works, without prior authorization from Virginia Tech. At the end of the one year period, either you or Virginia Tech may request an automatic extension for one additional year. At the end of the one year secure period (or its extension, if such is requested), the work will be handled under option 1 above, unless you request option 2 or 4 in writing.

___ 4. Release the entire work for Virginia Tech access only, while at the same time releasing the following parts of the work only (e.g., because other parts relate to publications) for worldwide access (check all that apply or provide an attached list):

___ Abstract and key bibliographic data (e.g., from submission form)

___ Files named as follows (e.g., separate PDF or multimedia files, as for chapters or figures):

Part B (use only if you checked 2 or 4 above). Our preference regarding being contacted to see if we will give written approval to expand the access to the above mentioned document is: (*choose one*)

___ in 1 year

___ in 3 years

___ probably never (e.g., since a publisher will release a book version soon)

Part C (optional proxy). To cover cases such as when one or more of the student and committee signing this form becomes inaccessible, each of the following people (indicated by their names printed)

Printed name of proxy: _____

Printed name of proxy: _____

Printed name of proxy: _____

is authorized to serve as a proxy in submitting future versions of this form, so submittals with any of these proxies signing are officially recognized just as if the student and full committee signed.

The above mentioned document has been reviewed and accepted by the student's advisory committee. The undersigned agree to abide by the statements above, and agree that this Approval Form updates any and all previous Approval Forms submitted heretofore.

Signed: _____

(student)

(date)

Committee: _____

(committee member)

(date)

DRAFT: Letter to Virginia Tech Students Preparing an ETD

Dear Student Preparing an ETD,

Your electronic thesis or dissertation (ETD) will contribute to worldwide graduate education as we build a Networked Digital Library of Theses and Dissertations (NDLTD) in collaboration with other scholarly institutions. We are writing to address concerns and questions you may have about how this relates to other types of publication. Please read the questions and answers that appear below, and feel free to contact us at etd@vt.edu if you have further questions.

The Electronic Thesis and Dissertation Initiative at Virginia Tech has several goals, including to help you in your career, to help other learners and researchers, and to make available many works that are now "lost" (e.g., theses that led to no other publications and that only are available through inter-library loan).

Most publishers contacted by the ETD Project Team support our initiative, realizing that theses and dissertations are very different from previously published or derivative books and articles. We believe that making ETDs available will supplement the efforts of publishers so both activities can proceed in harmony. We urge you to prepare your ETD to harmonize with publishing practices, and so that your research becomes as widely disseminated as possible, as soon as possible.

Thank you for your contributions to the Networked Digital Library of Theses and Dissertations and the Virginia Tech Electronic Thesis and Dissertation Initiative.

Sincerely,

Edward A. Fox, Professor
(for the Virginia Tech ETD Initiative)

Note: This letter represents views of the project team, not official University policy, but has been reviewed by a number of campus officials, outside experts on copyright, and publishers.

Questions and Answers

1. Why must I submit electronically?

By preparing an ETD and submitting it electronically you learn about electronic document preparation and about digital libraries. These skills will help prepare you for your future role in the Information Age, whether you teach, research, or use the research results of others.

Furthermore, you may be able to better convey the message of your thesis or dissertation in an electronic as opposed to a paper document. Thus, you can easily have color diagrams, color images, hypertext links, and even include audio, video, animations, spreadsheets, databases, simulations, virtual reality worlds, etc. in your appendices.

By submitting electronically you also allow your university to fulfill more economically its responsibilities of recording and archiving your thesis or dissertation. This is a key responsibility of the university, that is easier and less costly (in this time of tight budgets) to fulfill when the work flow involves electronic documents.

Note that electronic submission is totally separate from electronic access. So, please realize that regardless of what is arranged in terms of access to your work, electronic submission is required, unless special circumstances arise and are accepted by the

Graduate School.

2. Why is there no paper version required any longer for University Libraries?

The University only requires electronic submission (though your committee and/or department may still want a paper version). It commits to electronic archiving of works received, making sure that these will be accessible in the future, regardless of changes in media and standards. This is a firm guarantee so you need not worry. Furthermore, paper documents can easily be produced from electronic documents, but not vice versa.

By not accepting paper, the University reduces handling and library costs, saves you money, and makes it possible for access to increase.

3. How will people be able to access my ETD?

If you allow your ETD to be freely available worldwide, which we recommend (see below for reasons, and for discussion of other options), we will work to make your ETD as easily available as possible. First, we will allow access over the WWW, so people can link to our collection for browsing, and even link directly to your ETD (with a special type of URL that is not subject to change). Second, in the record for your ETD that will be in the Virginia Tech library catalog, we will have link information, so those searching that catalog can link directly to the ETD. Third, we will provide one or more search "engines" so that people can search the Virginia Tech ETD collection using "full-text" searching. Fourth, we will have a mechanism so that your ETD can be found by any seeking to search the NDLTD (i.e., the full distributed collection of ETDs made available by institutions that are part of the initiative). Fifth, we will work with 3rd party organizations, such as UMI and OCLC (a not-for-profit in Dublin, Ohio that provides library cataloging and other services to libraries), to encourage them to provide access as well as archiving services.

4. Why should I make my ETD freely available?

The world of scholarship depends on people making their research available to others. When that is done electronically, more people can get access at lower cost, and more knowledge transfer occurs. This can stimulate education and research. It also can ensure that many people give credit to you for your work, and that your research is cited in others' publications, which adds to your prestige and can help your future advancement. We can log all accesses and provide a report to you of the count, to pass on to your supervisors, if you request this.

Before theses and dissertations were available electronically, not many were read. Electronic access multiplies the number of times works are read by a factor of ten or more. Since you spent a great deal of time on your research, it should encourage you to know that others are reading that work. Your literature review may guide others, and your results may save others the time of redoing your study.

With electronic theses and dissertations, students and universities can more easily share knowledge, with much lower costs. We believe that about 200,000 theses or dissertations are completed each year. It would greatly aid graduate education if as many as possible of these were made freely available.

Since we aim to maximize access, which seems especially appropriate for a land grant university, we will not charge and so will not have any royalties to share.

5. What are the options regarding electronic access?

Virginia Tech gives you four options regarding affording electronic access to your ETD. The [Virginia Tech Electronic Submission Approval Form](#), that must be signed by you and your committee when you turn in the your final work, indicates your choice.

1. The first option, which we recommend (see above) is to make it freely available worldwide. This option may be viewed as *publishing* the work, so that book or journal publishers may see a conflict with their publishing part or all of your thesis or dissertation.
2. The second option, which we hope will be needed rarely, is to make the ETD freely available only to computers at Virginia Tech. This option addresses situations such as when a book is planned, and the book publisher worries that prior *publishing* through a digital library will hurt sales.
3. The third option is to secure the work, even disallowing access to computers at Virginia Tech. This option addresses situations such as when a patent application is planned, or when proprietary interests are at stake. In such cases, Virginia Tech may have some rights to related intellectual property, and so does not want you to release the work without its (Virginia Tech Intellectual Properties) permission. However, this option can only apply for one year (or two, if an extension is requested), whereupon another one of the options must be selected (else the first option will be assumed chosen).
4. The fourth option is to release the entire work at Virginia Tech (as in option 2 above) and also to release parts (which you identify by listing names of files --- so you need to divide up your submission in parts, such as with one chapter per file) for worldwide access. This option addresses situations where, say, a chapter relates closely to a journal article, so access to that chapter needs to be restricted, while at the same time the rest of the work does not and will not related to

You should consider these options carefully. Feel free to ask Virginia Tech Intellectual Property or Legal Counsel for advice, and to discuss this with your advisor. If you intend to work with a publisher regarding journal or book publications, be sure you understand their policies and any agreements you would sign.

We are happy to explain these options further, beyond what

appears below, and are actively working to document publishers' views regarding ETDs. Please note that you can help us prepare guidance for other students by completing the questionnaire we provide about reasons for your option selection, and about student practices and plans regarding publications related to theses or dissertations.

6. What if I want to write a book related to my thesis or dissertation?

We realize that some students, especially in the humanities, prepare books related to their theses or dissertations. In those cases we strongly recommend that you approach likely presses, and ask for advice regarding what level of access to allow to your work.

Usually, books that relate to theses or dissertations turn out to be significantly changed as part of the editorial process. This makes it likely that those interested in your work will buy your book when it comes out, even if they have reviewed your ETD.

In some cases it appears that electronic release of early versions of a book leads to greater sales of such books. Indeed, having an electronic work made available on the Internet, and telling a publisher that there have been a large number of electronic accesses to that work, may help you land a book contract.

However, since publishers vary widely in their policies, it may be wise to share this letter and other documents about the ETD initiative with publishers to which you are likely to submit your work. For example, if they consider the market of your work to be very small, they may not want to publish it if it has been made accessible worldwide.

We are open to discussions with publishers regarding policies or helping in the publicity process (in case they do publish your work). For example, we could create a bibliography or database of ETDs and list all related articles, books, or other publications

for each.

7. When should I be contacted about changing status of access?

Part B of the form provides 3 time periods as options so the graduate school knows when to re-contact you to see if there might be an increase in access allowed to your ETD. We recommend 1 year as the most suitable choice.

For various reasons, the situation that exists when you first prepare your ETD may be altered with the passage of time. For example, if a book has been published, and gone out of print, and the publisher has no plans for further printing, you may change status from a restriction to campus, to allow worldwide access.

Similarly, if one part of your ETD was held back from worldwide access because it relates closely to a journal article, and that journal article has appeared in print, the journal publisher is likely to be willing to allow you to release your ETD with the missing part included. That is, if the publisher wanted right of first publication, once widespread distribution of the journal has taken place, that publisher may give you written permission to broaden access to your ETD.

8. How does the ETD project relate to UMI?

UMI is a corporation in Ann Arbor, Michigan that maintains a microform archive of about 1.5 million dissertations, as well as an online service called Dissertation Abstracts. Most dissertations written in the US are submitted to UMI for archiving on microfilm, from which microform or paper copies can be produced. UMI functions as an on-demand book publisher that eliminates the editorial process. One of the services they offer is to help you regarding copyright and working with publishers.

They accept electronic submissions as well as paper submissions. The latter are scanned in and OCR'ed, but in most

cases current technology does not yield as good a result as would come from an electronic submission. UMI plans to make available online electronic versions of all works they receive after 1996.

Few masters theses are sent to UMI. The ETD initiative aims to handle the hundreds of thousands of theses that UMI does not receive each year.

UMI has a representative on the Steering Committee and on the Technical Advisory Committee for the NDLTD. More information about UMI can be found at <http://www.umi.com>.

The NDLTD project focuses on graduate education and raising the level of knowledge transfer. Since students may wish to read a thesis or dissertation that was prepared many years before, it is imperative that the NDLTD arrange for archiving of ETDs, so they can be accessed even when media and technology change. UMI, as well as OCLC, is interested in providing such archival services.

9. What do I need to know about signing agreements with publishers?

When you have your research published in a conference, book, or journal, you usually sign some type of agreement with the publisher. You should read that agreement carefully before signing, making sure you understand AND AGREE with the terms and conditions. If you don't, you may want to change the agreement in connection with discussion/negotiation with the publisher, and possibly with advice of legal or other counsel. The agreement should be explicit about what future rights of use you retain. If you want to include the materials in a dissertation or to reuse the materials for teaching or a book chapter, say so.

As the author you are entitled to discuss your plans with the publisher. We encourage you to obtain an agreement that allows you to include your research in a freely available electronic thesis

or dissertation.

During these negotiations you may want to discuss matters of timing and revision. You have the right to negotiate with a publisher to reduce access to your ETD to your university only for a limited amount of time, if they request this as a condition on publishing your article. However, most publishers consider a thesis or dissertation to be quite different from a journal article. Typically the article is much shorter than the chapter or full work, has been revised as a result of the editorial process and peer review, and sometimes has several authors, resulting in many publishers having no concern regarding fully accessible ETDs.

10. What if I want to have a journal article as a chapter in my thesis?

If you have published an article or articles before you turn in your thesis or dissertation, and you wish credit for that for your graduate requirements, you have a number of options. These should be discussed with your committee, and possibly with your publisher. First, you can simply cite that publication in your references. Second, if the publisher has the publication online, you can link or point to it (with permission of the publisher, who usually has protection so that paying customers or subscribers are the only ones allowed access). Third, if the publisher gives you a signed release, you can include the publication in your thesis or dissertation as allowed in that release. If the publisher restricts access in that release, say to your university, you may want to have 2 versions of your thesis or dissertation--one with and one without the chapter (e.g., published article) in question.

This matter may be avoided if your thesis or dissertation talks about your research in a very different way from the published article. That often makes sense, since articles are typically short, and your thesis or dissertation may be the only place where all the details, data, tables, and other aspects of your research are made available.

Remember that preparing a thesis or dissertation is part of your graduate experience, one aim of which is to prepare you to be a part of the world of research and publication. While this letter may help make the philosophy of the ETD Initiative clearer, the ultimate success of the initiative depends on you. We hope you will treat this a part of your educational experience, and will take steps when you deal with publishers to help other students gain the widest possible access to your research.

[\[For Students\]](#) [\[Table of Contents\]](#) [\[FAQ\]](#)

ETD Initiative

Questions? etd@vt.edu

Last revision: April 18, 1997



Electronic Thesis and Dissertation Project

Acceptable File Formats

Here is the current list of file formats we accept thus far.

Thesis Body

- PDF
- DVI
- ETD-ML

Text

- ASCII (.txt)
 - SGML
- Note: We recommend Unicode for non-Roman characters.

Images

- PDF (.pdf)
 - use Type I PostScript fonts
 - JPEG (.jpg)
 - CompuServe GIF (.gif)
 - TIFF following version 6.0 or later, including CCITT G4 (.tif)
 - CGM Computer Graphics Metafile (.cgm)
 - PhotoCD
- Note: We recommend a minimum of 600 dpi resolution for images of pages with text.

Video

- MPEG (i.e., MPEG-1, MPEG-2) (.mpg)
- QuickTime - Apple (.mov)
- Audio Video Interleaved - Microsoft (.avi)

Audio

- MPEG-2
- CD-DA
- CD-ROM/XA (A or B or C)
- AIF (.aif)
- SND (.snd)
- WAV (.wav)
- MIDI (with timing information) (.midi)

Authoring

- Authorware
- Director (MMM, PICS)

Special

- Spreadsheet - Excel (.xcl)
- AutoCAD (.dxf)

Referring

- "handles" as URNs (URLs don't last very long)
- ISBN, ISSN



[Back to Students](#)

[\[For Scholars\]](#) [\[For Students\]](#) [\[For Universities\]](#)
[\[Sponsorship\]](#) [\[Table Of Contents\]](#) [\[Search\]](#) [\[FAQ\]](#)
