ACM Multimedia '96
Course MA 2

TUTORIAL TITLE: Building and Applying Digital Libraries I

INSTRUCTORS: Edward A. Fox, Robert M. Akscyn

SCHEDULE: Monday, Nov. 18, 1996, 9 a.m. - 12:30 p.m.

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

COURSE DESCRIPTION:

Objectives:
Attendees should be able to help in design, development, evaluation, and standardization efforts related to digital libraries. They should understand the key aspects of representative digital library projects, as well as their successes, failures, and implications for the future. They should become familiar with the underlying technologies for digital libraries, such as information retrieval, hypertext and electronic publishing.

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).

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
    + OCLC [link] http://www.oclc.org:5047/oclc/research/projects/core/
  - TULIP [link] http://www.elsevier.nl/info/projects/tulip.html
    + 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://wwwqbic.almaden.ibm.com/
    + rights management [link] http://204.146.47.71:80/is/dig-lib/dlfnct4.htm
* Enhance WWW (hypertext):
  - HyperWave [link] http://www.tu-graz.ac.at/0x811b0205_0x00071882;sk=D40F3456
  - HyperWave server features[link]
  - HyperWave authoring features[link]
  - HyperWave authoring specs [link]
  - Harmony orientation [link]
  - Harmony screens [link] http://ei.cs.vt.edu/~cs5604/Adv/Adv-Harmony.html
  - Harmony information structuring [link]
  - Harmony viewers [link]
* 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 - 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])
  - Courses (Interactive Learning with a Digital Library in CS) [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
      + Virtual Computer History Museum
        http://ei.cs.vt.edu/CSNotes-cgi/infoBase/infoBase.
        pl?showPage+./001.sub/012.sub
      + Timeline ([link])
* DLI [link] http://www.grainger.uiuc.edu/dli/national.htm
  - CMU (Carnegie Mellon U.)
    + Informedia [link] http://www.informedia.cs.cmu.edu/research/index.html
    + NetBill [link]
      http://www.ini.cmu.edu/NETBILL/publications/CompCon_TOC.html
  - 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://walrus.stanford.edu/~testbed/
    + Today's DLs [link]
      http://Walrus.Stanford.EDU/~testbed/testbed.slides/P002.htmls
    + Infobus [link] http://Walrus.Stanford.EDU/~testbed/testbed.slides/P003.htmls
    + COS - before[link] http://Walrus.Stanford.EDU/~testbed/cos/slides/P011.html
    + COS [link] http://Walrus.Stanford.EDU/~testbed/cos/slides/P002.html
    + COS - services list [link]
      http://Walrus.Stanford.EDU/~testbed/cos/slides/P015.html
    + COS event services ([link])
      http://Walrus.Stanford.EDU/~testbed/cos/slides/P023.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
  - Standards [link] http://etd.vt.edu/etd/faq/formats.html
  - Principles [link] http://etd.vt.edu/etd/principles.html

RECOMMENDATIONS:

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

ACKNOWLEDGEMENTS

Sponsors and Partners: Elsevier, IBM, NSF, OCLC, SURA
Co-PIs, Students, Staff
Projects
* NSF IRI-9116991, CDA 9312611, EID-9109853, CDA-9308259
* NCSTRL (ARPA, NSF)

Introduction

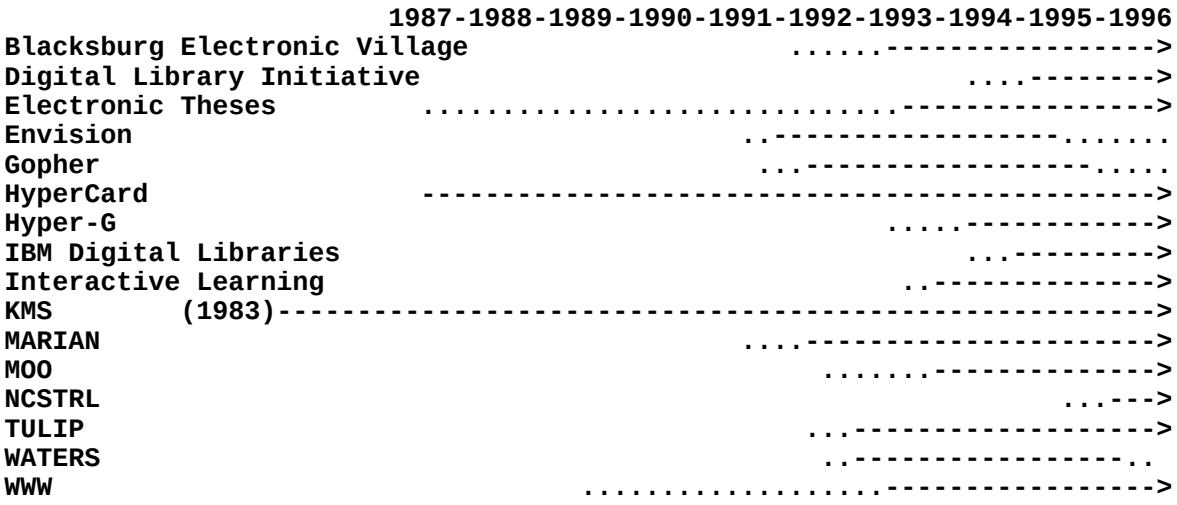


Figure 1: Timeline of Recent Information and Digital Library Systems

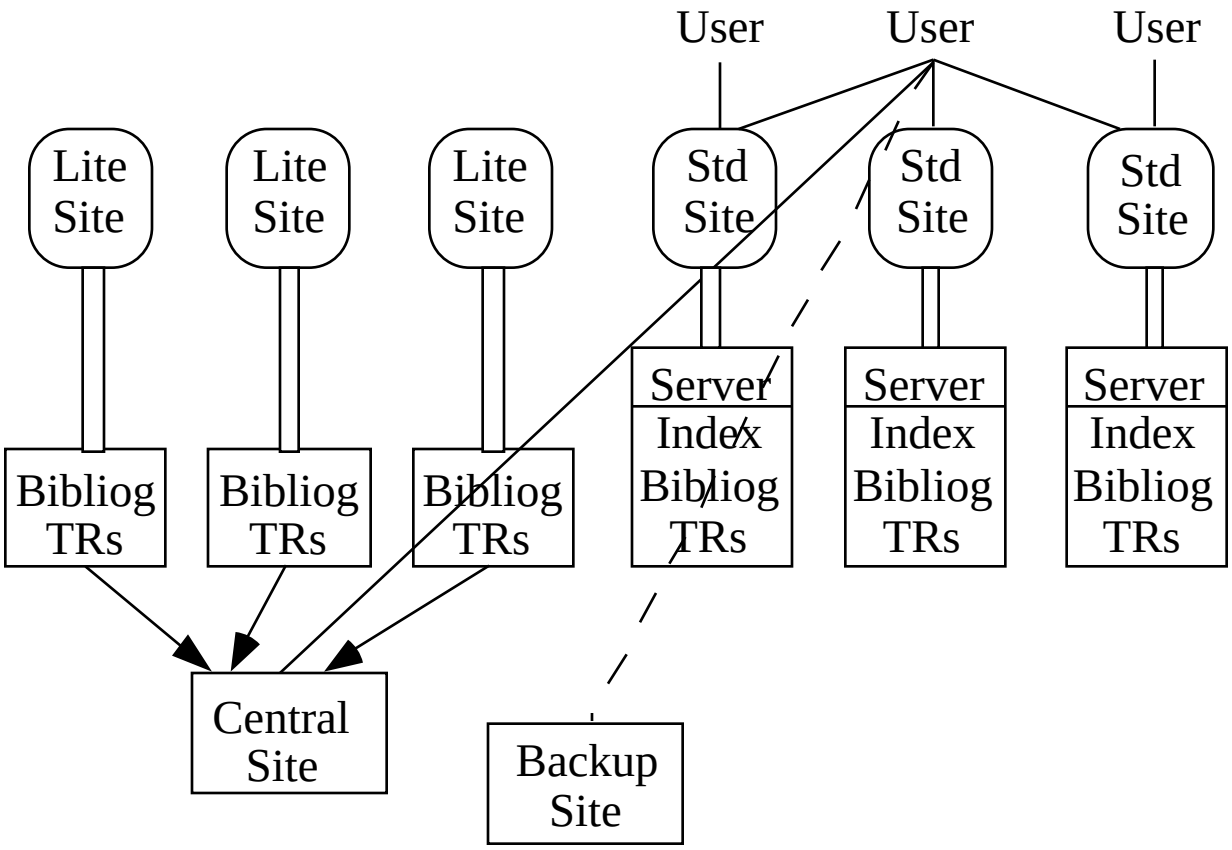


Figure 2: NCSTRL Architecture

Netlib Repository at UTK and ORNL

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

There have been 11,785,446 requests to this repository as of Sat Jul 13 02:23:01 EDT 1996 .

Software, papers, etc.

- Browse the Netlib repository
- Search the Netlib repository

Services provided at Netlib

- Conferences Database
- National High-Performance Software Exchange (NHSE)
- Numerical Analysis Net (NA-Net)
- Performance Database Server
- Top500 Supercomputer Sites

Information about Netlib

- Frequently Asked Questions about Netlib (FAQ)
- Netlib Editors
- Netlib Mirror Sites
- Netlib Server Statistics

Netlib Maintainers

Netlib Attribute-Value Database

This is an index of the Netlib Attribute-Value Database. Please type a query in the search dialog. You may use freeWAIS-sf query syntax.

This is a searchable index. Enter search expression:

Submit

Attribute 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 seach 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

The CORE Project: Overview

The CORE project is an electronic library prototype the 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" formating 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.

TULIP The University Licensing Program

When you scroll further down this page you'll find

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

Digital Library is the Information Age tool for extracting value from your information assets in new ways.

Digital Library Adapts To Your Needs

Every digital library presents unique challenges. That's why IBM Digital Library is not only comprehensive but scalable and flexible. To see digital libraries that exist today, go directly to the Case Histories. Quick and thorough overviews of IBM Digital Library solutions for organizations in government, higher education, media and cultural institutions are included. For fast access use the button below the Wheel.

On the other hand, get the essentials on building your own digital library by taking five steps through the main components of IBM Digital Library. The Wheel is your guide. Jump in anywhere, and explore the key functions: Create & Capture, Search & Access, Distribution, Rights Management, and Storage & Management.

To have an IBM representative with digital library expertise bring you detailed information for your own Digital Library project, use the More Info link at the bottom of every page. And help us serve you better by taking a few minutes to complete our survey while you're at it.

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 (Government, Higher Education, Media, 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.

Government

Government agencies collect and store massive amounts of information. Widespread bureaucracies need to share the information assets they command. Meanwhile, the public's desire to gain access to public information seems endless.

IBM Digital Library provides the foundation for linking hubs across a range of government sites. A user needn't travel to the capital to find governmental information. IBM Digital Library enables people to find what they need from wherever they are.

Derwent
With IBM Digital Library, Derwent offers network access to U.S. patent information. Users see a Lotus Notes interface while they access 20 years of patent data with weekly updates. A full patent document can be ordered from Derwent with a single button click.

Higher Education

The concept of a school without walls has been discussed for over a century. By setting the stage for distance learning and access to learning archives, both using multiple forms of media, IBM Digital Library finally makes the "open classroom" truly open. All the important documents can be available on-line 24 hours a day, 7 days a week.

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.

Media
The convergence of entertainment, information and technology is dynamically changing the media and entertainment industries. IBM Digital Library brings intelligence and value to content creation and hosting, asset banking, workgroup access, electronic commerce, rights management and protection, royalties payments and licensing, distribution and archiving. What's more, IBM Digital Library can be "under the covers," whereby content owners can establish their own branded image for the services and archives they create.

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.

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.

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.

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.

The QBIC Project

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

Query By Image Content(QBIC)

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 complex queries of large image databases. The queries are based on image content -- color percentages, color layout, textures, and shapes of image and their objects. Some of this technology is currently available in IBM's Ultimedia Manager product.

To try the World Wide Web QBIC search engine on a database of approximately 1,900 images, click on the button below:

Try our latest browser: Try Demo

You must have a HTML3.0 capable browser to run the above demo. If you don't, try downloading the new IBM WebExplorer for OS/2 or Netscape.

Otherwise, you can use our older browser: Try Demo

Did You Enjoy Our Demo?

If you did, we are looking for early-adopters in selected application areas. If you have a web site with images and would like to use this technology, please contact The QBIC Group

If you didn't, we would like to know if we can improve it in some way. If you have any suggestions or comments please mail to 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:

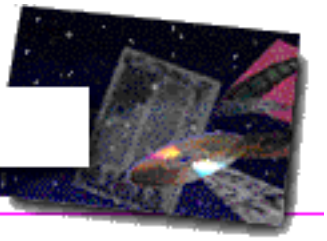
Your e-mail address:

Other links related to QBIC:

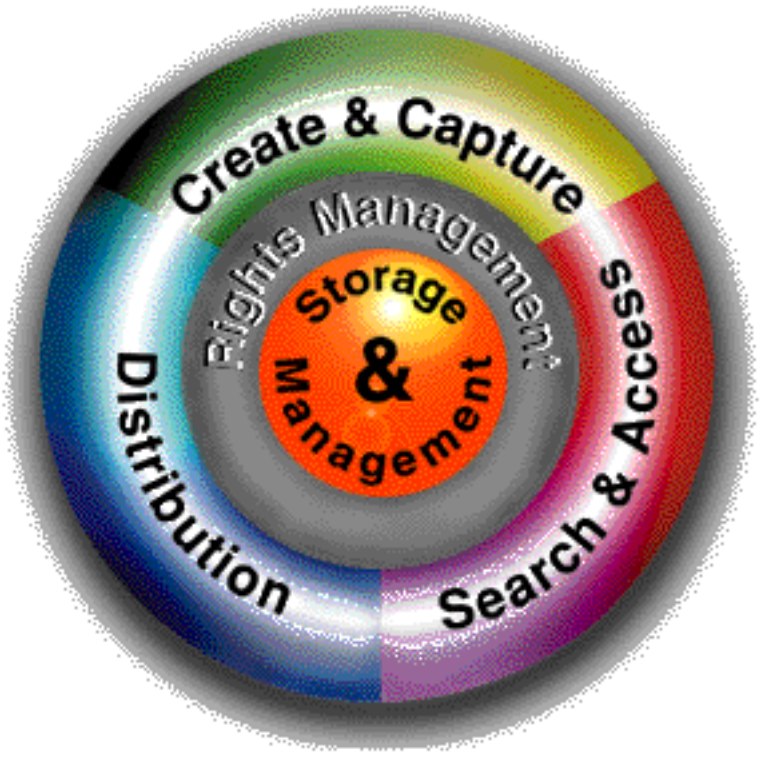
- A show and tell of the QBIC technology.
- Technical paper requests on QBIC.
- Ultimedia Manager 1.1 - A product that incorporates some of the QBIC technology.
- DB2 Extenders, which will soon include QBIC technology.

IBM

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 digitzed 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.

User: www-anonymous

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 AUTHORING 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]

More attribute information.
Author: iicm
Parent(s):

IFABO 95
About IICM, HMS & their projects
IICM's Public Services
About HyperG
Working with HyperG in Paderborn

User: www-anonymous

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 browers 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.

User: www-anonymous

KEY FEATURES

HYPERMEDIA AUTHORING

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.

More attribute information.
Author: gmesaric
created: 96/04/16 09:24:32
modified: 96/04/17 10:28:14

User: www-anonymous

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)

User: www-anonymous

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.

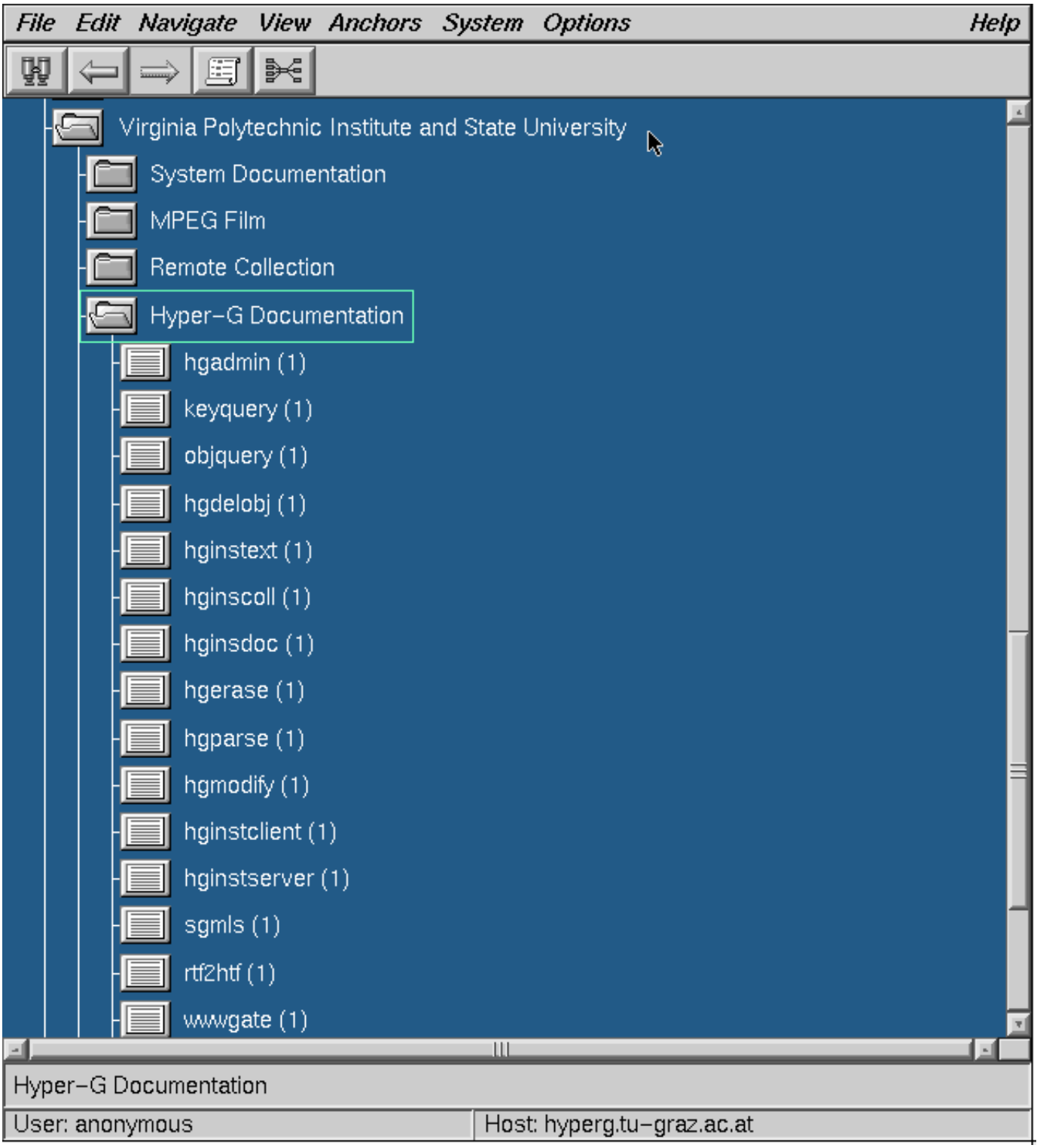
3D Information Landscape

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

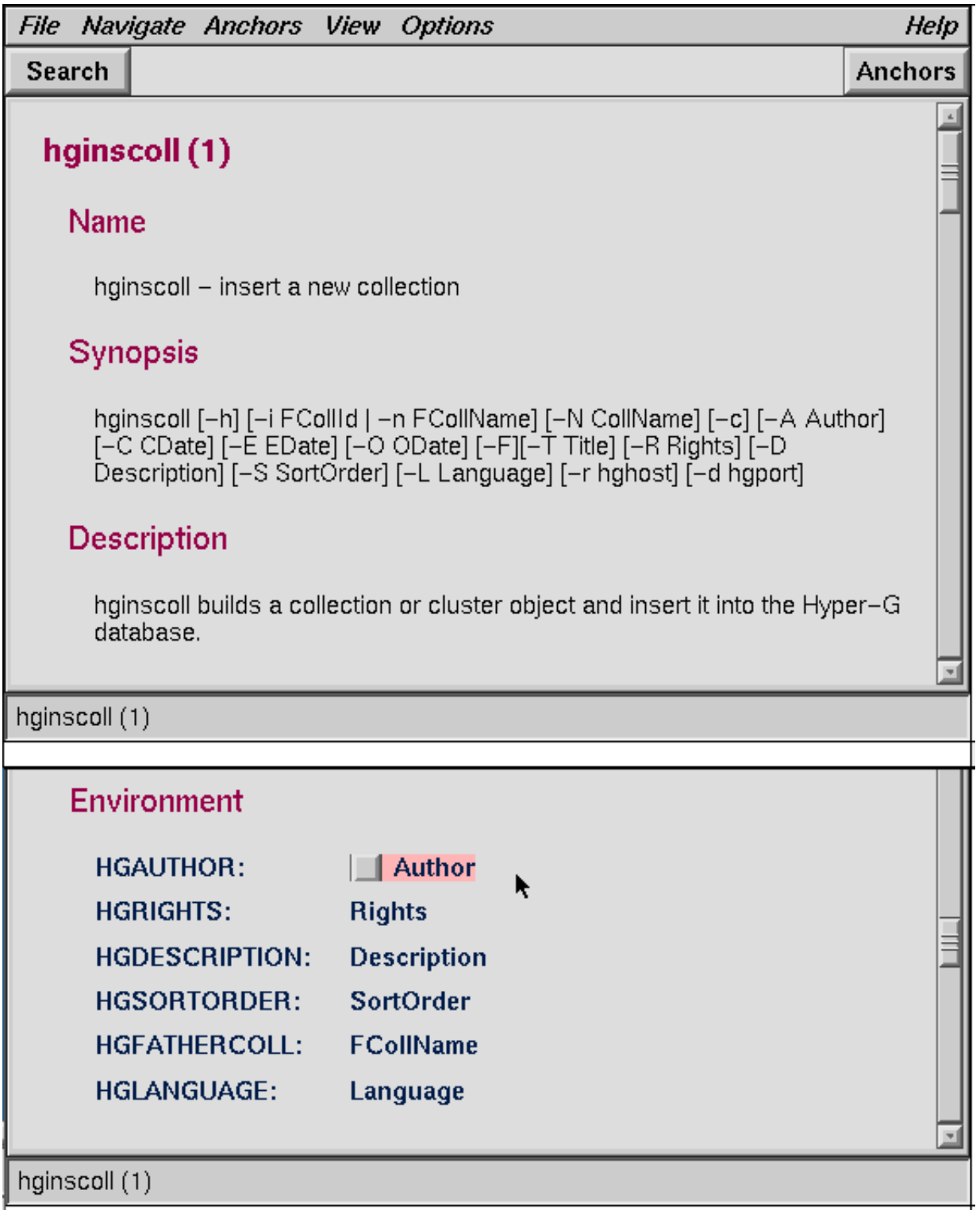
More attribute information.
Author: iicm
created: 95/12/18 08:19:00
modified: 95/12/20 08:06:04

Parent(s):

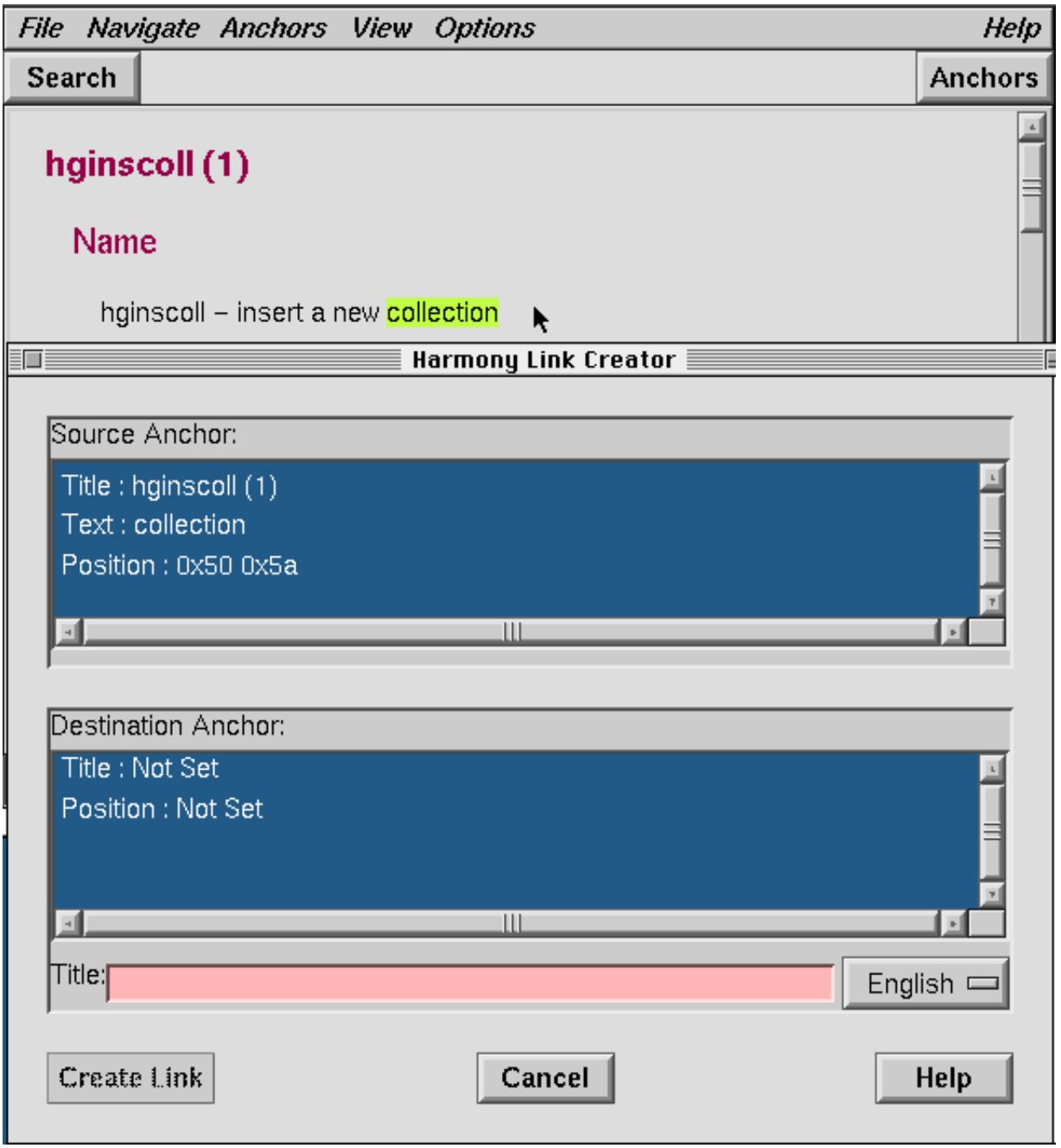
Harmony's orientational aids



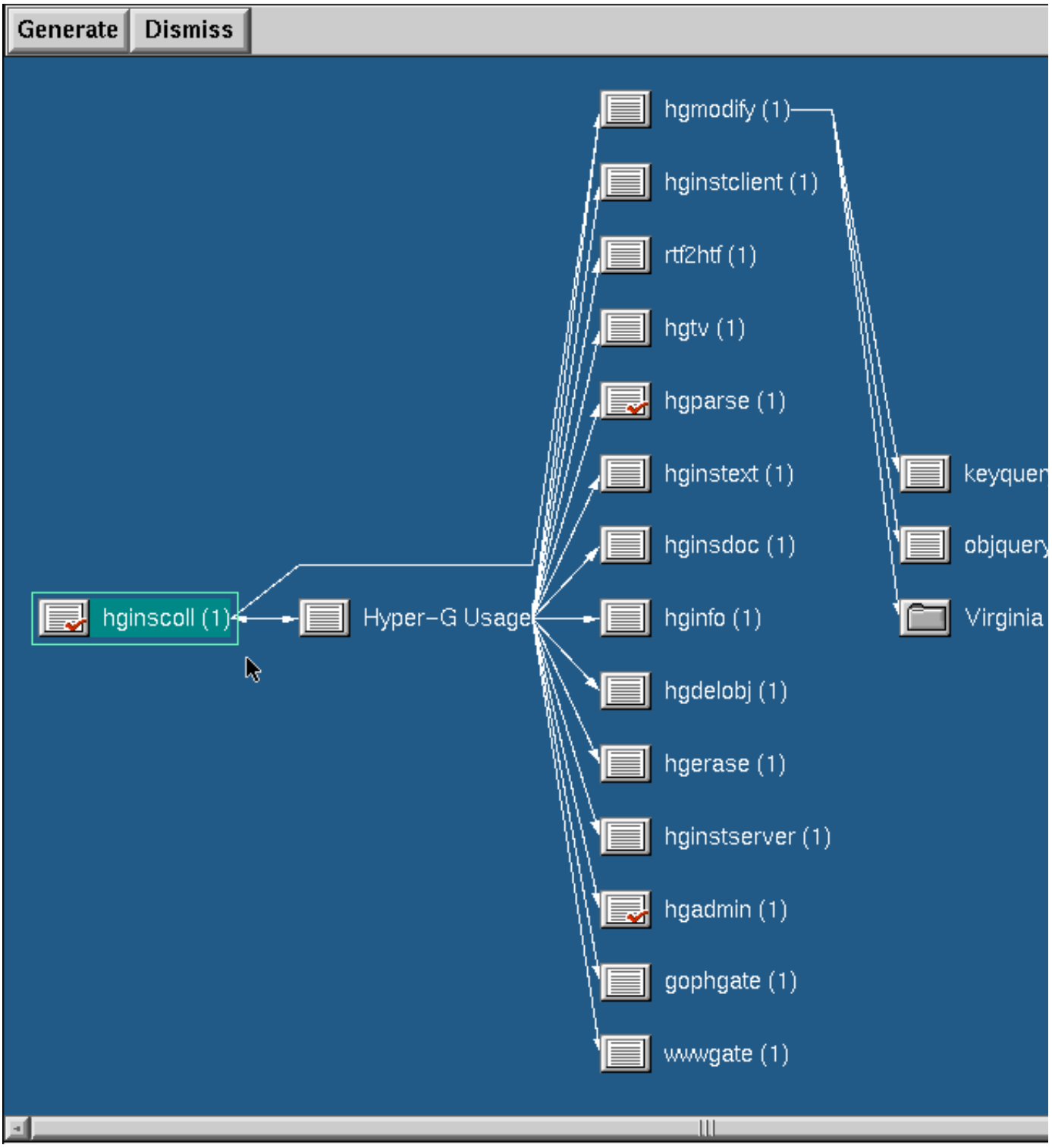
- viewing two text nodes and marking an anchor



- making a link



- viewing a local map


Generate
Dismiss
hgmodify (1)
hginstclient (1)
rtf2htf (1)
hgtv (1)
hgparse (1)
hginstext (1)
hginsdoc (1)
hginfo (1)
hgdelobj (1)
hgerase (1)
hginstserver (1)
hgadmin (1)
gophgate (1)
wwwgate (1)
keyquer
objquery
Virginia
hginscoll (1)
Hyper-G Usage

User: www-anonymous

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.

More attribute information.
Author: iicm
created: 95/12/18 07:32:25
modified: 96/01/17 14:06:58

Parent(s):

Information structuring in Harmony

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.

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

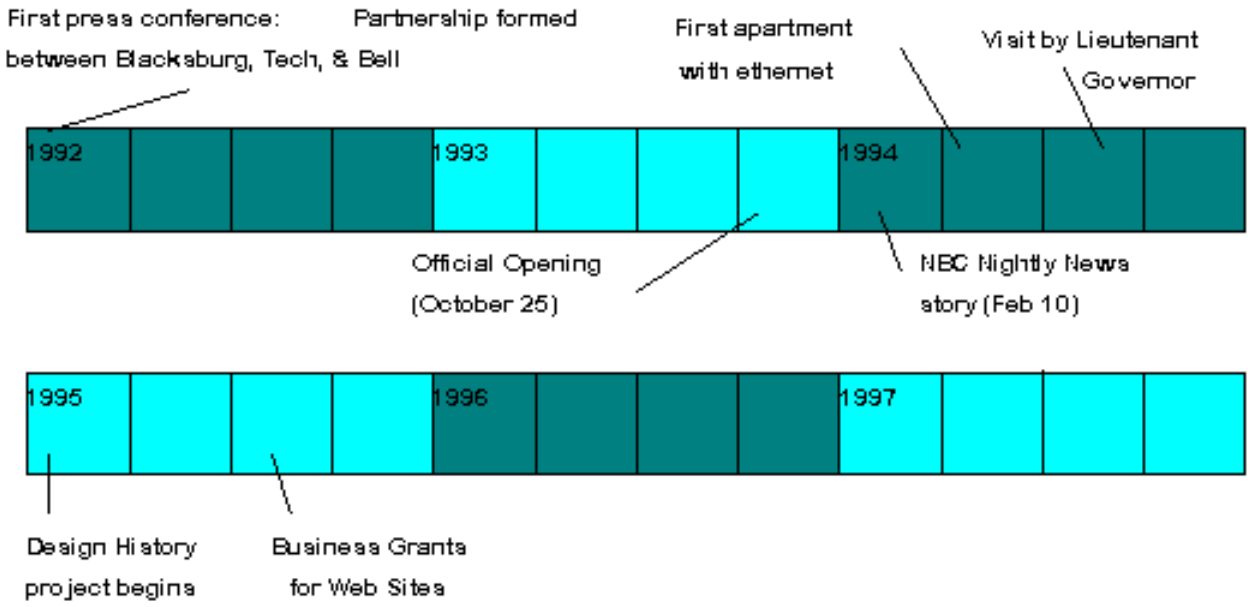
This project is supported by NSF Grant CDA-9424506. A copy of the grant proposal is online.

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: 1992 1993 1994 1995 1996 1997

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

HistoryBase Main Page | Main Timeline | Quarter Year View | Contribute | What's New? | Search | BEV Homepage

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

HistoryBase Main Page | Main Timeline | Quarter Year View | Contribute | What's New? | Search | BEV Homepage

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]

NCSTRL
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
hi-resolution tiff image
inline gif image

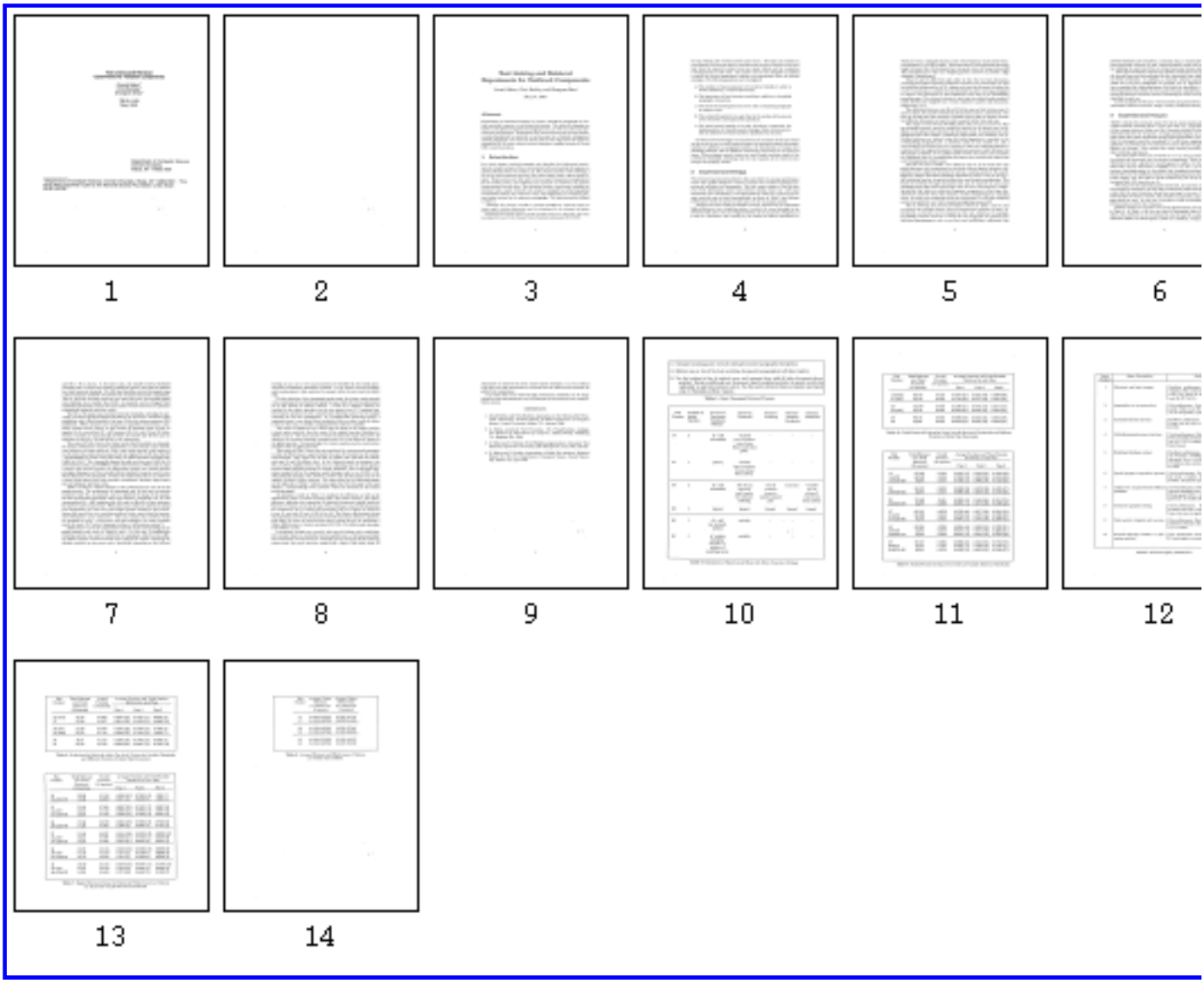
1 Display page

- 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.

[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]

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. T 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 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 Inquery
U. Mass. CIIR demo of Inquery 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
NCSTRL
Networked Computer Science Technical Report Library
Univ. of Wales Cardiff CS Courseware
Courseware on Algorithms, AI, C, Graphics, Image Processing, Parallel Processing, Vision, X

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 Courses

Welcome to one of the largest (over 25 courses, over 4500 nodes) 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

- 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 Programming in C
- 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
- 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
- CS4504: Computer Organization
- CS4624: Multimedia, Hypertext and Information Access
- CS5014: Research Methods in Computer Science

- CS5024: Models and Analysis
- CS5034: Models of Computation
- CS5114: Theory of Algorithms
- CS5204: Operating Systems
- 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

Searching All Courses and other Pages on ei.cs.vt.edu

Summary about Harvest collection from ei.cs.vt.edu

Usage Statistics

See also

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

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

Pointers to Previous Years' Materials

Usage Statistics

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

Calendar for CS5604

- **1. DL:** *Introduction, Digital Libraries*
  FOXE88d, 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], 10/31
- **11. HT + MM:** *HT + Multimedia*
  FOXE91b, WALL91
  11/5, 11/7
- **12. MM:** *Multimedia*
  PHIL91a, GREE92
  11/12, 11/14
- **13. KB:** *Knowledge-Based Information Retrieval*
  FURN87, LEBO88, MALO87a
  11/19, 11/21
- **14. Break**
- **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

CS4624: Multimedia, Hypertext, Information Access
Table of Contents (Spring 1996)

Please read: News / Announcements
(Most recent update is for 960503.)
DEADLINE FOR TURNING IN WORK WAS MIDNIGHT !

Outline

Student Information

- **Home Pages for Tuesday** and **Thursday**
- **Progress Report**

Activities

- **Lab Sessions**
- **Lecture Notes**
- **Projects**
- **Quizes, Final (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**
- **Index**
- **Koofers (old quizes, final)**
- **Link Sets**
- **Tutorial on Multimedia (1994)**
- **Readings and References**
- **Searching All CS Class Notes Online with Harvest**

Syllabus (all together, as of 2/24/96, plus Outline) or in sections:

- **Calendar**
- **Department and Class Policies**
- **Instructor and GTA**
- **Syllabus Details: Grading, etc.**

Fall 1995 Version

Comment Form --- Email Submission

Note: "4984" is the number used 1995-96 but "4624" is the number to be used thereafter.

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 ; WWW: MOO Information
  - *Systems/Demonstrations:* Author/Editor, KMS, MIME, Mosaic, 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:* 3, 12; WWW: Virtual Reality
  - *Systems/Demonstrations:* AuthorWare, Director, ENVISION, Virtual Reality
- CR: Capture and Representation (3.5 wks)
  - Course Notes: 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:* 4, 5, 11, 14; WWW: CD-Video 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:* 6, 7; 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:* 10; WWW: Hyper-G, ILDLCS, OpenDoc (and Bento), QuickTime, ScriptX
  - *Systems/Demonstrations/Exercises:* ENVISION, HyperCard/Toolbook, Hyper-G, HyperTies, MARIAN, MHEG, ScriptX, WAIS
- 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:* 9, 15; book: Hodges+ Ch 20; WWW: Information on Videoconferencing
  - *Systems/Demonstrations/Exercises:* CU-SeeMe, MBone
- Post-Test and Review

Virtual Computer History Museum

Our group will design a prototype virtual museum of computer history on the World Wide Web. We envision various methods of accessing the 'exhibits' within the museum, such as:

- **a text-based chronology of events**
- **a graphical timeline of events**
- **a searching mechanism**
- **an image gallery**
- **pre-defined tours based on specific times or people**
- **and links to other computer-related museums**

The Virtual Computer History Museum Group consists of:

- Charles Atwood (atwoodc@vt.edu)
- Jon-Erik V. Lido (jlido@vt.edu)
- Mike Marston (marston@csugrad.cs.vt.edu)
- Andy Wagliardo (awagliar@core-dump.com)

Much assistance, guidance, support, and direction was provided by

- Dr. Edward A. Fox (fox@vt.edu)
- Dr. J.A.N. Lee (janlee@vtopus.cs.vt.edu)

of the Computer Science Department at Virginia Polytechnic Institute and State University (Virginia Tech).

Meeting place: McBryde 110, Derring 2069

Meeting time: Monday 2:00pm - 3:00pm, Thursday 12:15pm - 1:45pm

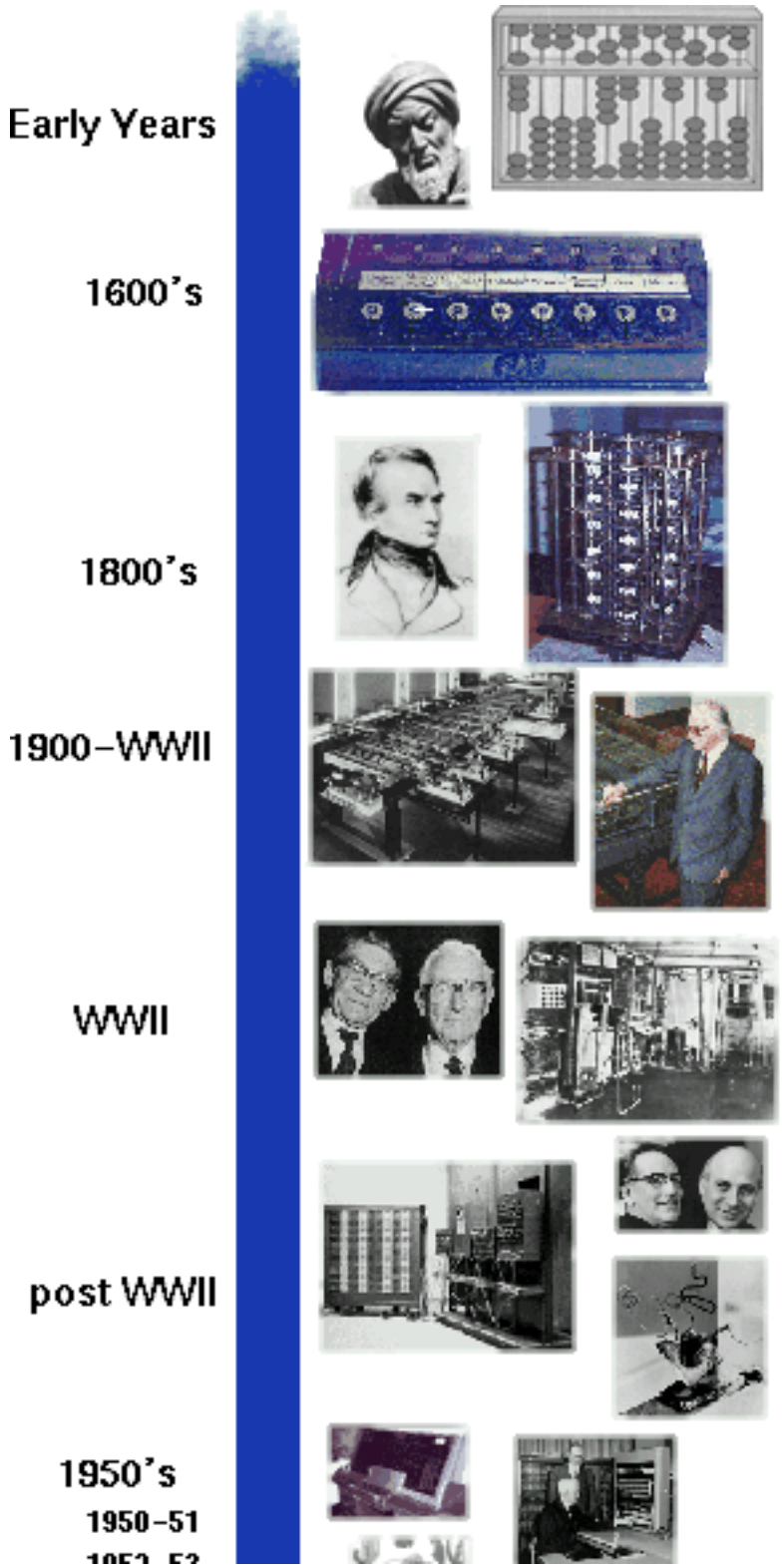
Contact: Andy Wagliardo ((540)232-3627 museum@mail.core-dump.com)

Mailing Address:
Andy Wagliardo
202 Barringer Hall
Blacksburg, Virginia 24060-0001

WWW Address:
http://ei.cs.vt.edu/CSNotes-classes/VirtualComputerHistoryMuseum

Timeline of Events in Computer History

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



DIGITAL LIBRARY INITIATIVE

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.

Funded through a joint initiative of the NSF/ARPA/NASA Digital Library Initiative

IEEE *Computer's* May 1996 Special Issue on Digital Libraries

Carnegie-Mellon University

Full-content search and retrieval of video

Principal Investigator: Howard Wactlar

Contact: Colleen Everet, (412)268-7674

Stanford University

Interoperation mechanisms among heterogeneous services

Principal Investigator: Hector Garcia-Molina

Contact: Maryanne Siroker, (415)723-0872

University of California at Berkeley

Work-centered digital information services

Principal Investigator: Robert Wilensky

Contact: Crystal Williams, (510)642-0930

University of California at Santa Barbara

Spatially-referenced map information

Principal Investigator: Terrence R. Smith

Contact: Patty Towne, (805)893-7665

University of Illinois at Urbana-Champaign

Federating repositories of scientific literature

Principal Investigator: Bruce Schatz

Contact: Susan Harum, (217)244-8984

University of Michigan

Intelligent agents for information location

Principal Investigator: Daniel Atkins

Contact: Laurie Crum, (313)763-6035

Outline
Sponsors
Links
Search
Home

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

Informedia project publications

Research Area Descriptions

- Image Understanding
- Natural Language Processing
- Speech Recognition
- Human Computer Interfaces
- Networked Data Transport
- Network Billing and Security

The NetBill Overview

NetBill: An Internet Commerce System Optimized for Network Delivered Services

Marvin Sirbu, Engineering and Public Policy Department
J. D. Tygar, Computer Science Department

Carnegie Mellon University
Pittsburgh, Pennsylvania 15213

As the explosive growth of the Internet continues, more people rely on networks for timely information. However, since most information on the Internet today is free, intellectual property owners have little incentive to make valuable information accessible through the network. There are many potential providers who could sell information on the Internet and many potential customers for that information. What is missing is an electronic commerce mechanism that links the merchants and the customers.

NetBill is a business model, set of protocols, and software implementation allowing customers to pay owners and retailers of information. While NetBill will enable a market economy in information, we still expect that there will be an active exchange of free information.

The highlights of the NetBill model include:

- Has a very low transaction costs for micropayments (around 1 cent for a 10 cents item)
- Protects the privacy of the transaction
- highly scalable
- certified delivery mechanism which delivers information goods if and only if the customer has payed for them.

This paper discusses the design of the NetBill protocol and our World Wide Web (WWW) prototype implementation

Note: the paper is contained in a single file (49Kbytes). You can also get a postscript version. The following links can be used to go to a specific section of the paper.

Next Up Previous
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.

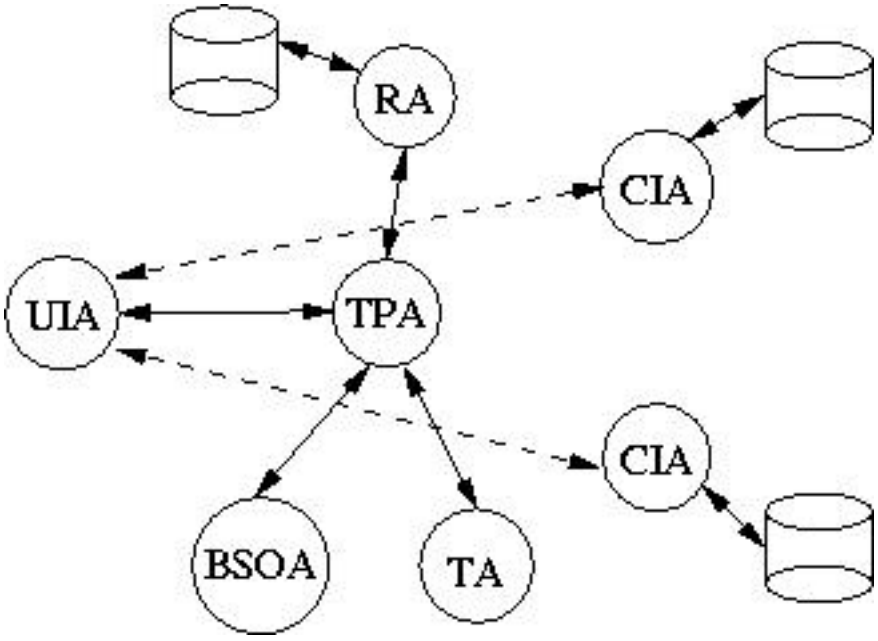


Figure 1: Schematic view of the UMDL architecture showing the agents in it and the communications between them. The solid lines represent the first part of the search, where the UIA looks for applicable collections. The dashed lines are the second part, when the UIA queries the collections that were 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

UMDL Organization

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

The above graph illustrates the 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 Digital Library Testbed Development

Department of Computer Science
Stanford University
Stanford, CA

Development team:

- Scott Hassan
- Andy Kacsmar
- Andreas Paepcke
- Tom Schirmer

What's new!

A major function of the Stanford Digital Library test bed is to allow experimentation with 'glue' for interactions with online services. In view of this requirement we chose a distributed object approach as our base technology. This decision was explained in the slides of our presentation to the first advisory board meeting on January 9, 1994.

We selected Xerox PARC's ILU as our implementation of distributed objects. ILU is roughly an implementation of the Common Object Request Broker (CORBA) standard, providing language bindings for C++, C, CommonLisp, Python and Modula-3. This means that we can do our implementations in any of these languages.

The best way to start is to look through the CORBA/ILU documentation, and then to look at some examples for the language you are interested in (see below). After you think you understand the basics (feel free to consult with us), ILU bugs can be reported to ilu-bugs@parc.xerox.com. General ILU questions can be directed to ilu@parc.xerox.com. To get on the ILU mailing list, send mail to ilu-requests@parc.xerox.com.

Note that for C++ users we have acquired a site license for the products ObjectCenter/ViewCenter/TestCenter. These are a debugger, motif interface builder and memory-leak-debugger/performance tuning tool respectively. They are available on HP platforms under /local/CenterLine/bin. For documentation, see Andreas.

- DL Object Interchange service and examples

We are working on a technical performance evaluation of ILU, HTTP, and basic TCP. We will be comparing ILU's performance to IBM's DSOM and maybe Microsoft's COM.

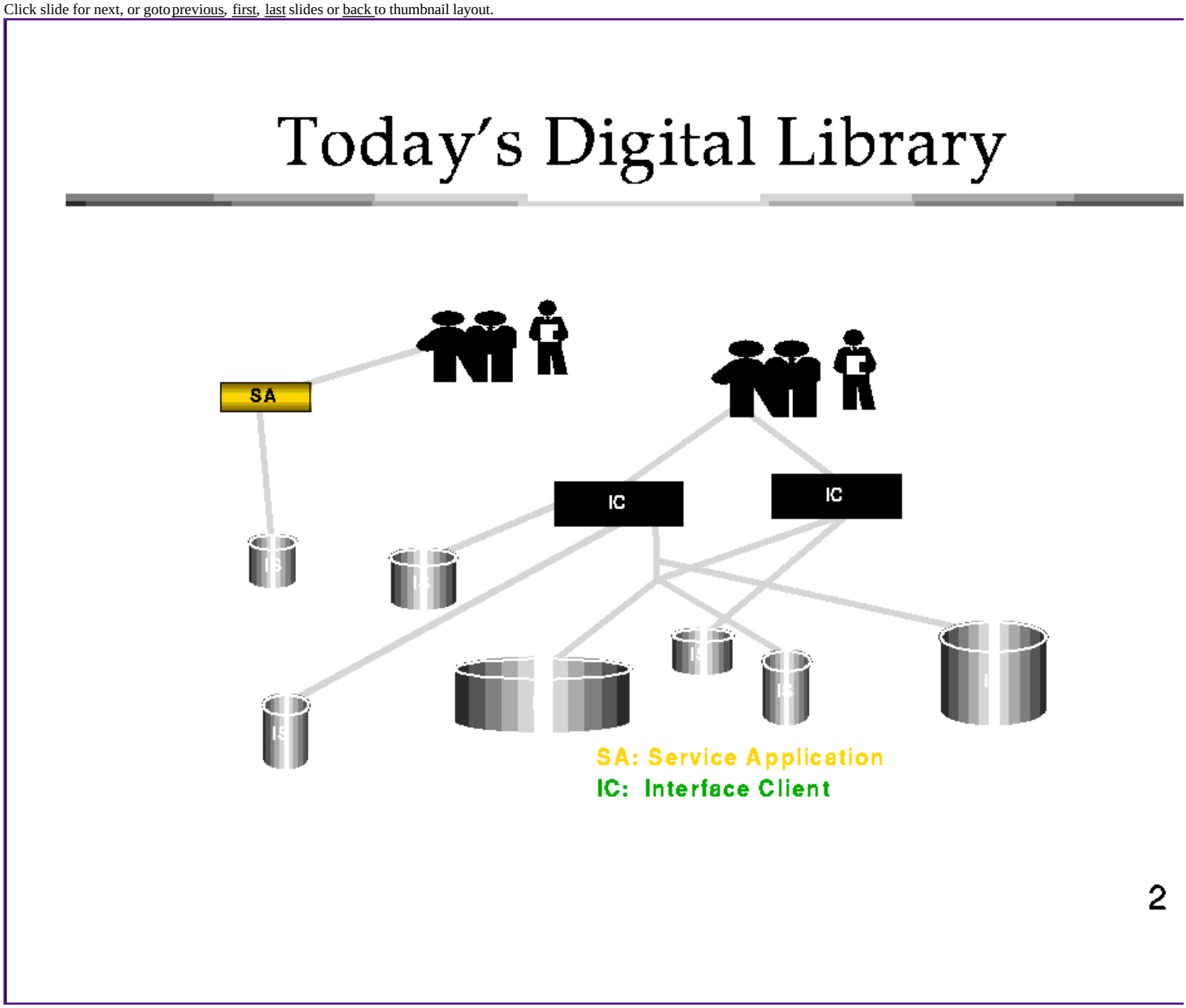
- ILU -- Our installation and examples.
- Python -- Our installation of Python programming language
- ObjectCenter -- Discussion of ObjectCenter
- COS -- Digital Library Testbed Common Object Services
- CVS -- Our use of CVS in the testbed.
- Various Manuals (CVS, Python, ILU)

- Testbed Activity Slides

Digital Libraries Webmaster
Webmaster@diglib.stanford.edu

Testbed: Slide 2 of 12.

Click slide for next, or goto previous, first, last slides or back to thumbnail layout.

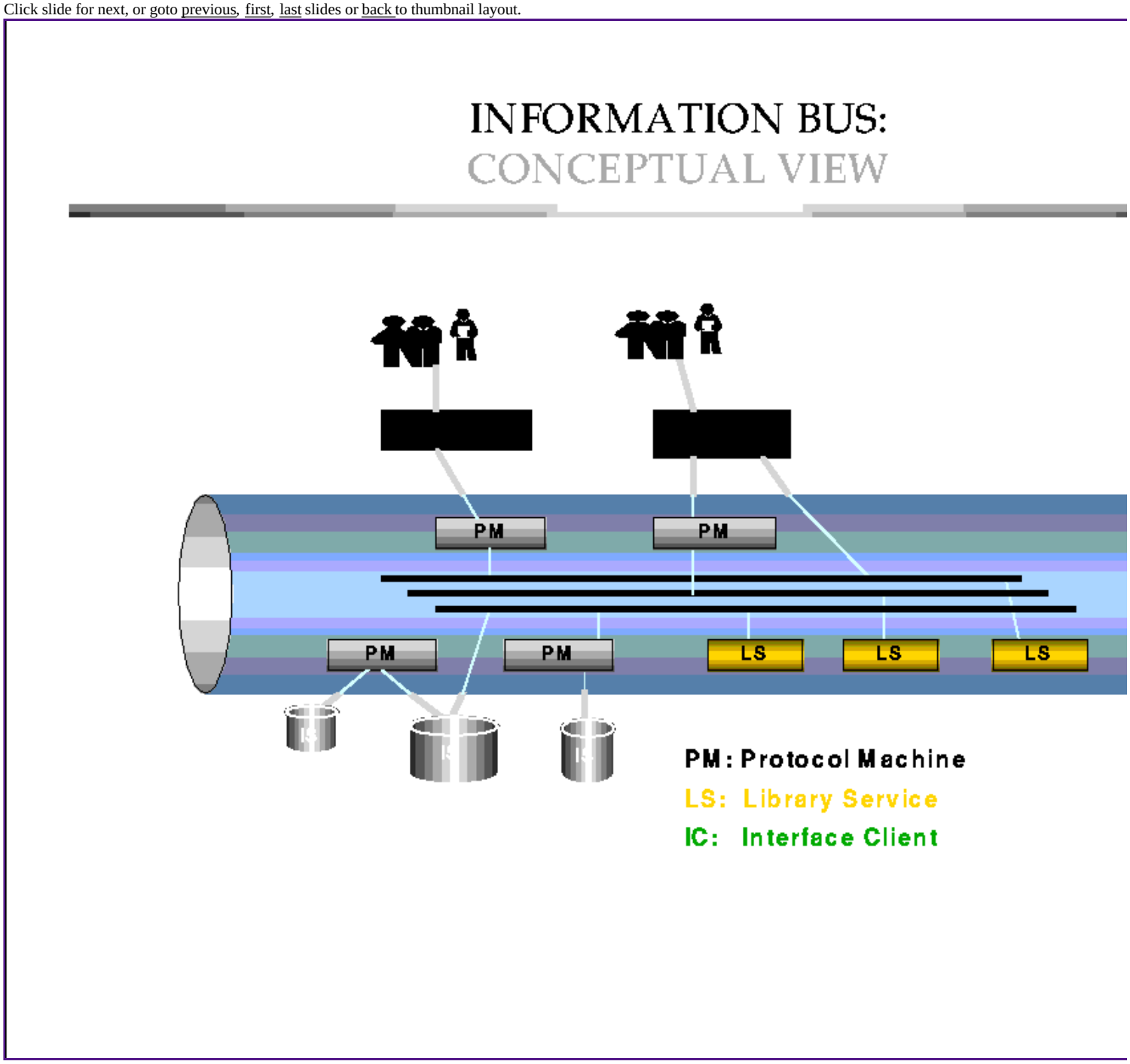


Click slide for next, or goto previous, or back to thumbnail layout.

Digital Libraries Webmaster
Webmaster@diglib.stanford.edu

Testbed: Slide 3 of 12.

Click slide for next, or goto previous, first, last slides or back to thumbnail layout.

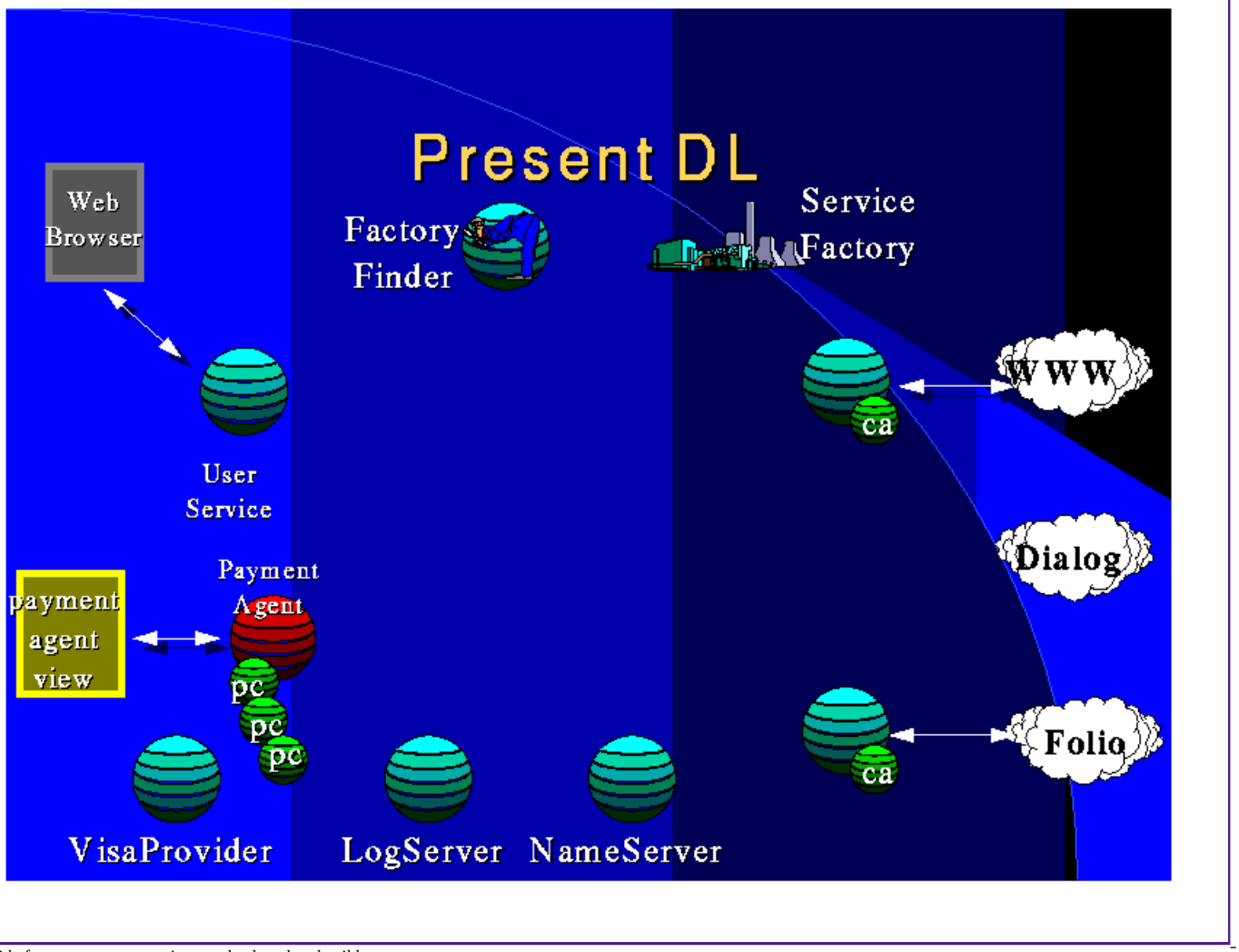


Click slide for next, or goto previous, or back to thumbnail layout.

Digital Libraries Webmaster
Webmaster@diglib.stanford.edu

COS - Common Object Services: Slide 11 of 49.

Click slide for next, or goto previous, first, last slides or back to thumbnail layout.

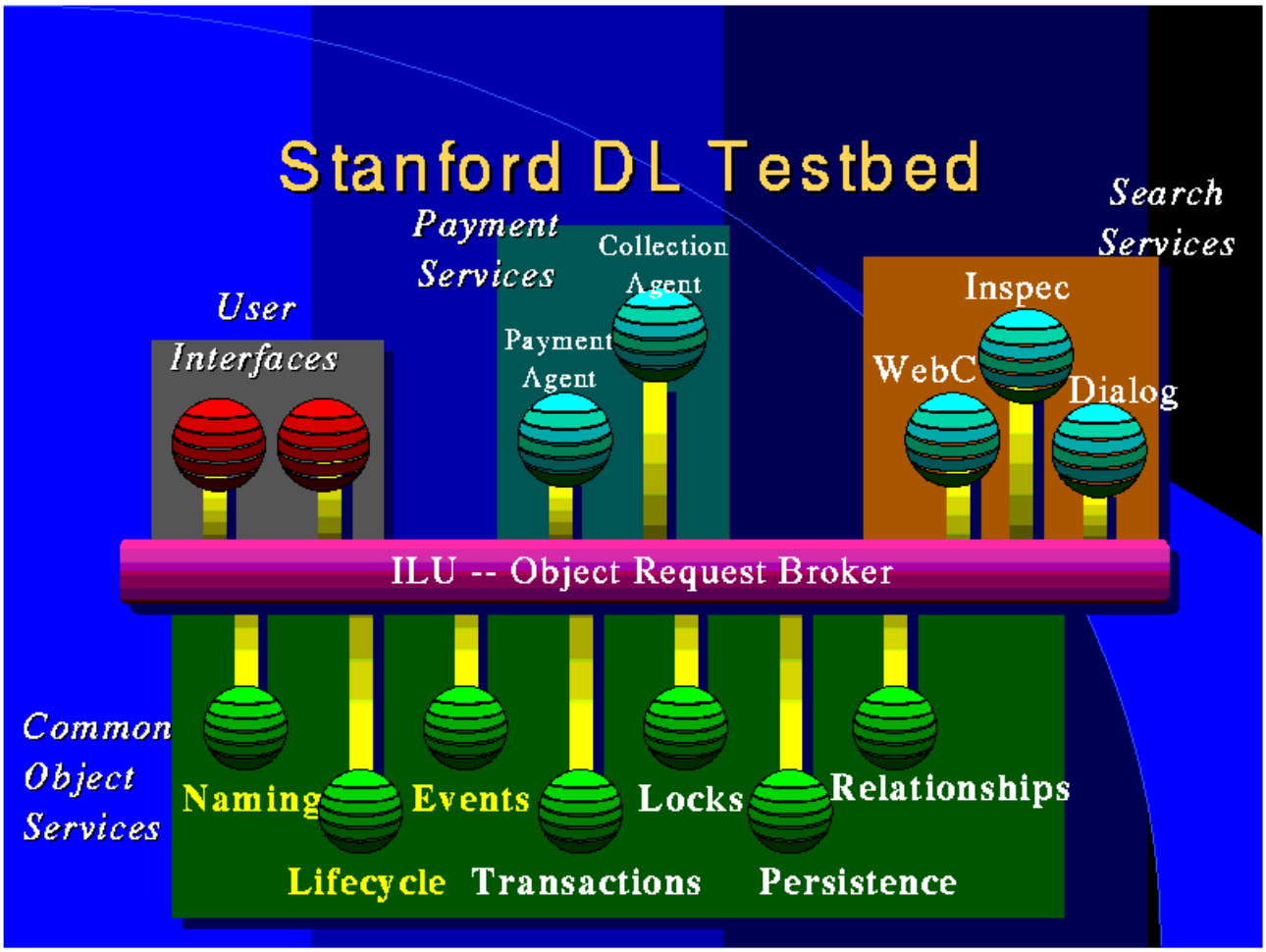


Click slide for next, or goto previous, or back to thumbnail layout.

Digital Libraries Webmaster
Webmaster@diglib.stanford.edu

COS - Common Object Services: Slide 2 of 49.

Click slide for next, or goto previous, first, last slides or back to thumbnail layout.



Click slide for next, or goto previous, or back to thumbnail layout.

Digital Libraries Webmaster
Webmaster@diglib.stanford.edu

COS - Common Object Services: Slide 15 of 49.

Click slide for next, or goto previous, first, last slides or back to thumbnail layout.

Click slide for next, or goto previous, or back to thumbnail layout.

Digital Libraries Webmaster
Webmaster@diglib.stanford.edu

COS - Common Object Services: Slide 23 of 49.

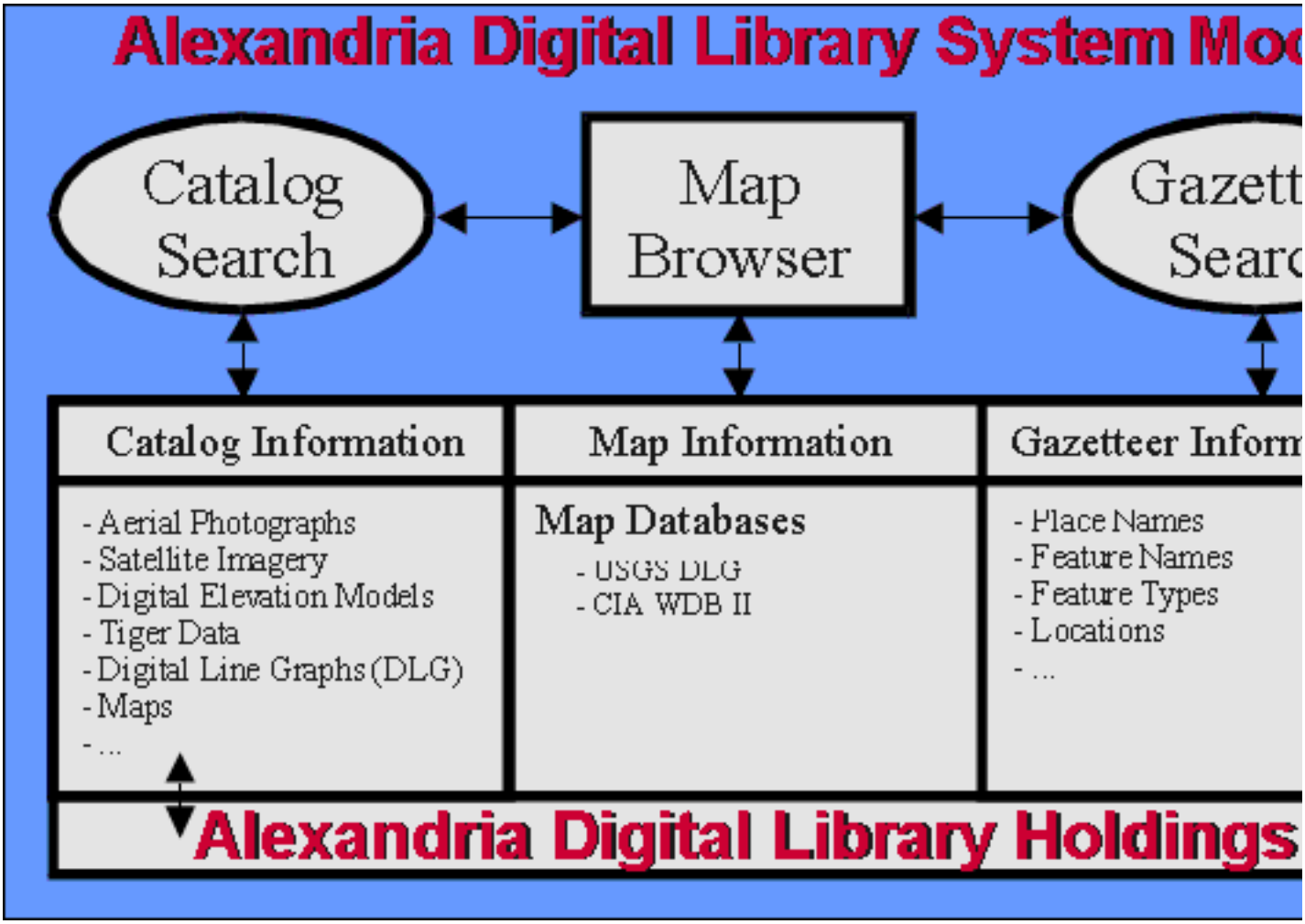
Click slide for next, or goto previous, first, last slides or back to thumbnail layout.

Click slide for next, or goto previous, or back to thumbnail layout.

Digital Libraries Webmaster
Webmaster@diglib.stanford.edu

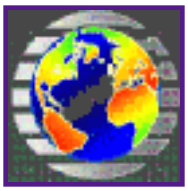
Tutorial Table of Contents

Prev Next



Conceptual model of the Alexandria Web interface

1996 ANNUAL REPORT Table of Contents



Alexandria Digital Library

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, 195 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 Repostories (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

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

<ETD>

Electronic Thesis and Dissertation Project

Local Information

- Library Collection of ETDs
- Recent ETD Additions
- Submission Guidelines
- Workshops
- Campus Computer Labs

General Information

- Frequently Asked Questions (FAQ)
- Press Releases and Project Descriptions
- Scenarios and Proposals
- Templates and Instructions for LaTeX

Other Project and Technical Information

Other Project Participants (Sites)

Quotes

"An Electronic Thesis or Dissertation (ETD) is a document that explicates the research of a graduate student -- it is expressed in a form simultaneously suitable for machine archives and

worldwide retrieval." -- Neill Kipp

"For students, the electronic thesis or dissertation is easier and cheaper to prepare, simpler to distribute, and more flexible in format. It encourages greater creativity and expressiveness on the part of the author by permitting inclusion of hypertext links and multimedia elements such as digital audio and video recordings." -- Dr. John Eaton

"We are embarked upon a national effort to train the future professorate, and the next generation of scholars, to be information literate, so that they can publish electronically and make effective use of digital libraries." -- Professor Edward Fox

Sponsorship

- Southeastern Universities Research Association (SURA) has funded the 1996 project: *Development and Beta Testing of the Monticello Electronic Library Thesis and Dissertation Program*
- The U.S. Department of Education's Fund for the Improvement of Post Secondary Education (FIPSE) program has funded a proposal entitled *Improving Graduate Education with a National Digital Library of Theses and Dissertations* to start 9/1/96. Authorized funding for the first year is in the amount $69,762. Anticipated future funding for years 2 and 3 are: $69,337 and $68,941. If all federal funding is received as planned, the total would be $208,040. Virginia Tech will provide institutional support in the amount of $182,417, which gives federal/nongovernmental percentages 53.3/46.7. Additional in-kind support from other partners is expected of over $130,000:
- Support for the FIPSE proposal has been promised by: ACM, Adobe, Council of Graduate Schools (CGS), Coalition for Networked Information (CNI), Cornell Digital Library Research Group, Council of Southern Graduate Schools (CSGS), IBM, OCLC, State Council of Higher Education for Virginia (SCHEV), Southeastern Library Network (SOLINET), SURA, UMI, University of Utah Graduate School

---

 **Back to *parent node***

<ETD> *ETD project*
etd@vt.edu

Last revision: July 31, 1996

Virginia Tech <ETD> Submission Form

This form will not work properly unless
you are using Netscape Navigator 2.0 or better.

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:
Select the type of document you are submitting.

- (•) Thesis
- ( ) Dissertation
- ( ) Special Report
- ( ) Major Paper
- ( ) Technical Report

Name:
Enter your name just as it appears on the title page.

Email address:
Enter your email address.

Adjust the width of your browser so that both ends of the "Title" field are visible.

Title:
Enter the title just as it appears on the title page.

Degree:
Enter degree.

Department:
Enter department.

Approved:
Enter the names and positions (ex: Chair, Co-Chair) of the people on your review committee.
Also enter the email address of the Chair, or the most senior Co-Chair.

	Chair	**email:**	

Keywords or phrases:
Enter keywords separated by commas(ex.: vortex field, CALS, DTD, DARPA).

Date of defense:
Enter date of defense (ex.: September 10, 1995).

Abstract:
Enter your abstract below. If you are cutting and pasting **be sure to read the help file section on abstracts first!** If you do not it is likely that your abstract will look very wrong when viewed by the Graduate School and when publicly available on the WWW.

Availability:
Select when you want your work to be available to the public.
- (•) As soon as possible.
- ( ) In 3 months (while preparing your manuscript for publication.)
- ( ) In a year (needed for clearance of sponsor proprietary rights and/or for patent application proceedings)

Copyright Statement:
I hereby grant to Virginia Polytechnic Institute and State University or its agents the right to archive and display my thesis or dissertation in whole or in part in the University Libraries in all forms of media, now or

hereafter known. I retain all proprietary rights, such as patent rights. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation. ◉ I agree

If you do not agree to the above statement please select "Clear/Erase" below and contact the Graduate School.

Signed Title Page:
I realize that my submission will not be processed until I give the Graduate School a **printed** copy of my title page that has been **signed** by the people on my review committee. ○ Yes ◉ No

---

Number of Files:
Please enter the number of files you are going to send to the Graduate School. 1

Preview:
Press "Preview" to continue. Preview

Scholarly Communications Project

Electronic Thesis and Dissertation Project

Frequently Asked Questions

about File Formats

What are the accepted file formats for ETDs?

Here is the current list of file formats we accept thus far. Please keep in mind that this list is growing.

Text

- ASCII (.txt)
- SGML according to the document type: "etd.dtd" (.etd)
  -- 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)

Electronic Thesis and Dissertation Project

Principles

Principle 1:

We should allow parts of ETDs to be encoded according to all widely used international standards (e.g., JPEG). Other representations should be selected based on agreement of an ETD Standards Committee.

Principle 2:

References to objects outside the ETD should be restricted to widely used reference methods that are descriptive in nature (e.g., give a page range in a proceedings of a conference whose city, dates, sponsor, name and editors are provided) or that follow some persistent naming scheme (e.g., ISBN, ISSN plus vol/no/page, URN).

Principle 3:

Encoding should be done in a way that allows recovery of all critical details. Thus, if a page of text is encoded as a bitmap, the smallest characters on the page must be clearly readable without any ambiguity. If an image is encoded, the smallest details relevant to the author's purpose in inclusion of the image must be accurately rendered. Thus, the emphasis should be on creating an archival representation, not on one that is easily or quickly rendered with current devices.

Principle 4:

Suitable metadata must be provided for all digital objects, as called for in ETD requirements or in the author's discipline, e.g.:

- **author/creator,**
- **permission details if not by the ETD author,**
- **date of origination if not that of the ETD,**
- **any details of origination and/or capture that would be needed by one wishing to correctly render the digital object, e.g.,**
  - **scanner make and model used,**
  - **settings/calibration at time of capture.**

Principle 5:

Quality is important. If analog devices are involved, they should be calibrated and tested in advance so an accurate recording is made.

Principle 6: