

Introduction to Digital Libraries



hussein suleman
uct cs honours 2005

Course Structure

- 16 lectures
 - intro
 - 11 dl topics - lectures
 - 3 dl issues – discussions
- 1 essay – discussion and critical analysis
- 1 programming assignment
 - some interesting information management service
- take-home final

Course Topics

- definitions and examples
- data/service model
- metadata, DC
- repositories, searching, software
- OAI, interoperability, architecture
- data encoding, compression
- IP/DRM, preservation, REST

Definitions and Examples



Example 1

The screenshot shows a Netscape browser window titled "Course: Introduction to Digital Libraries - Netscape". The address bar contains the URL <https://moodle-old.cs.uct.ac.za/course/view.php?id=14>. The page content is as follows:

- Page Title:** Introduction to Digital Libraries
- User:** You are logged in as [hussein suleman](#) (Logout)
- Course Path:** [CompSci](#) -> [CSC400DL](#) Turn editing on
- People:**
 - [Participants](#)
 - [Edit profile](#)
- Activities:**
 - [Assignments](#)
 - [Forums](#)
 - [Resources](#)
- Search:** Search forums
- Administration:**
- Topic outline:**
 - Course Tools
 - [News forum](#)
 - [General course discussion](#)
 - 1 Course Content
 - [Syllabus](#)
 - [Introduction \(pdf\)](#)
 - [Introduction \(ppt\)](#)
 - [XML and Metadata \(pdf\)](#)
 - [XML and Metadata \(ppt\)](#)
 - [OAI-PMH \(pdf\)](#)
 - [OAI-PMH \(ppt\)](#)
 - [Component Models \(pdf\)](#)
- Latest news:**
 - [Add a new topic...](#)
 - 14 Apr, 17:52 - hussein
Welcome to DL [more...](#)
- Recent activity:**
 - Your last login was:
Wed, 16 Feb 2005, 05:57 PM
 - Nothing new since your last login

The browser's status bar at the bottom shows the URL <https://moodle-old.cs.uct.ac.za/mod/resource/view.php?id=422>.

Example 2

The screenshot shows a Netscape browser window titled "UCT - Basic Search - Netscape". The address bar contains the URL: `http://alephprod.calico.ac.za:8991/F/?func=file&file_name=find-b&local_base=`. The page header includes the ALEPH logo and navigation links: Sign-in, End Session, My Library Card, Preferences, Databases, Feedback, Help, Browse, Search, Results List, Previous Searches, and Basket. Below the header, there are search options: Basic Search, Multi-field, Multi-base, Advanced Search, and CCL. The main section is titled "Basic Search" and contains the following form fields:

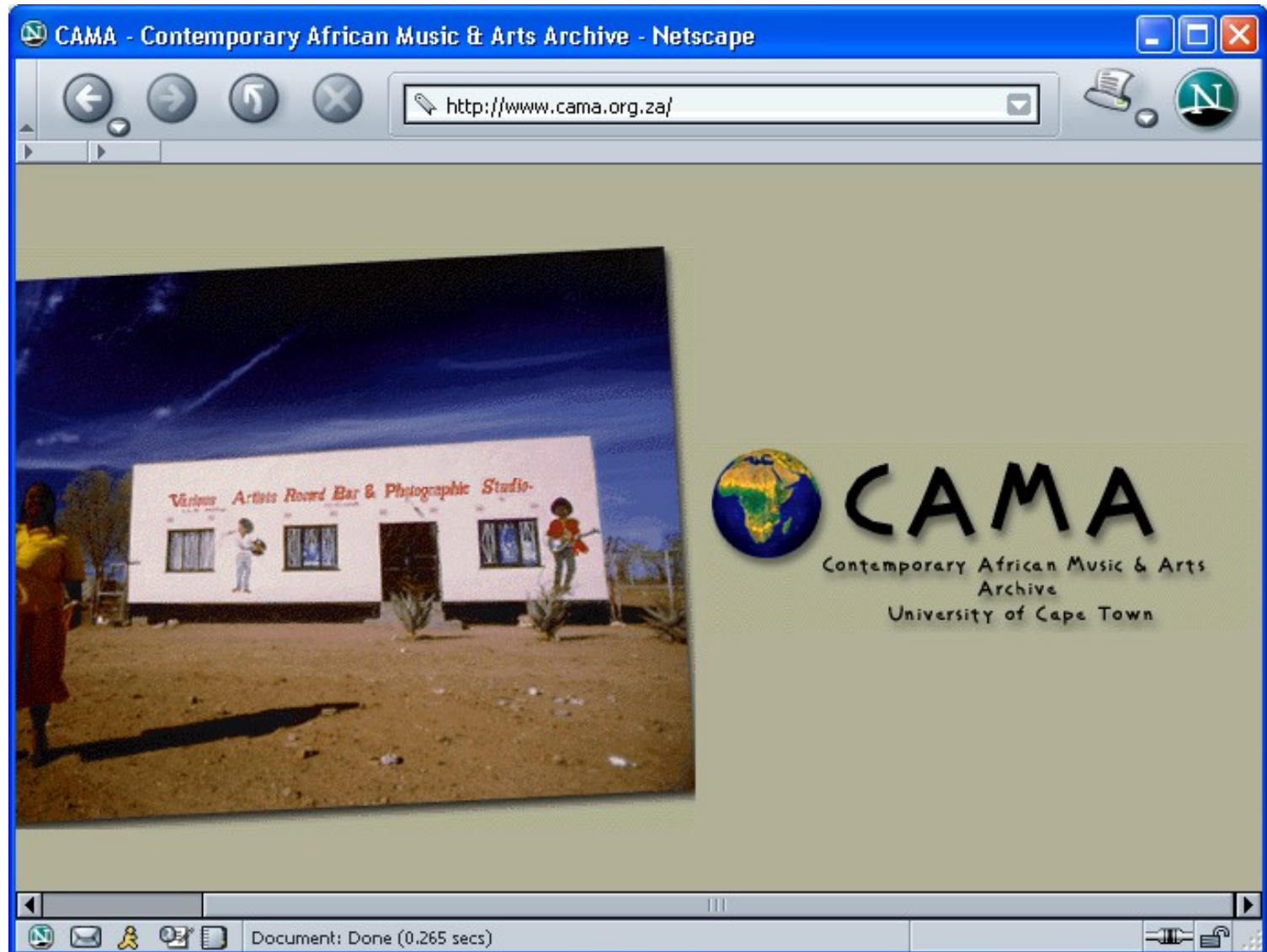
- Type word or phrase:
- Field to search: All Fields (dropdown)
- Words adjacent?: No Yes
- Base to search: University of Cape Town (dropdown)
- Buttons: Go, Clear

Below the search form, there is a "Limit search to:" section with the following fields:

- Language: all (dropdown)
- Year from: Year to: yyyy (Use ? for truncation when not using from/to)
- Format: all (dropdown)
- Location: All UCT Libraries (dropdown)

At the bottom, there is a "Search hints:" section. The browser's status bar at the bottom shows "Done" and various icons.

Example 3



Example 4

The image shows a screenshot of the Yahoo! homepage as it appeared in a Netscape browser window. The browser's address bar displays the URL <http://www.yahoo.com/>. The page features the large red "YAHOO!" logo in the center. Navigation icons for Personalize, Finance, Shop, Email, Messenger, and HotJobs are arranged around the logo. Below the logo, there is a section for "Yahoo! Personals" with links to "Find your Valentine", "Post a free ad", and "Benefit the American Heart Association". A search bar is located below the personals section, with a "Search" button and links to "advanced search" and "most popular". The page is organized into several columns of links and advertisements. On the left, there are sections for "New!" (linking to "Yahoo! Tax Center"), "Shop Find" (linking to "Auctions, Autos, Classifieds, Real Estate, Shopping, Travel, HotJobs/Careers, Maps, People Search, Personals, Yellow Pgs"), "Connect" (linking to "Chat, GeoCities, Greetings, Groups, Mail, Messenger, Mobile"), "Organize" (linking to "Addresses, Briefcase, Calendar, My Yahoo!, PayDirect, Photos"), "Fun" (linking to "Games, Horoscopes, Kids, Movies, Music, TV"), and "Info" (linking to "Finance, Health, News, Sports, Weather"). On the right, there is a "Yahoo! Travel Fare Sale!" advertisement with the text "Nationwide fares from \$73 - Last Day!" and links for "Flights - Hotels - Cruises - Vacations". Below the travel ad is a "Personal Assistant" section with a "Sign up" link. At the bottom right, there is an "In The News" section with a timestamp of 3:50pm and a list of news items: "Bush: 'Game is over' for Saddam Hussein", "NASA still considering foam launch damage", and "CIA officer killed in Afghan grenade". At the bottom left, there is a "Yahoo! Movies" section with "Now in Theaters" (listing "Deliver Us From Eva in 10 Days" and "Shanghai Knights") and "Coming Soon" (listing "Daredevil", "The Jungle Book 2", "The Life of David Gale", and "Gods and Generals"). A small movie poster for "Shanghai Knights" is displayed next to the "Coming Soon" list. The browser's status bar at the bottom shows "Document: Done (14.796 secs)".

Example 5

The screenshot shows the IMDb website in a Netscape browser window. The browser title is "The Internet Movie Database (IMDb). - Netscape" and the address bar shows "http://www.imdb.com/". The website features a navigation bar with links for "Register", "NOW PLAYING", "MOVIE/TV NEWS", "MY MOVIES", "FUN & GAMES", "MESSAGE BOARDS", "U.S. MOVIE SHOWTIMES", "HELP & GUIDE", and "IMDbPro" (with a "FREE TRIAL!" badge). Below the navigation bar is a section titled "Also Available" with links for "Top Movies", "Photo Galleries", "Video/DVD", "Browse IMDb", and "Independent Film".

The main content area is titled "The Internet Movie Database" and includes the text "Visited by over 13 million movie lovers each month!". Below this, there is a section for "IMDbPro.com" with a link to "Click here for a free trial." and a "Road to the Oscars" diamond-shaped logo.

On the left side, there is a search box with a dropdown menu set to "All" and a "Go!" button. Below the search box are links for "More searches", "Tips", and "IMDbPro.com free trial".

Below the search box is a section titled "Road to the Oscars" with a list of links: "Globe winners", "Critics list", "Awards news", "Photos", and "Polls".

Below the "Road to the Oscars" section is a section titled "Tops at the Box Office" with a list of movies: "1 The Recruit", "2 Final Destination 2", "3 Biker Boyz", "4 Kangaroo Jack", and "5 Darkness Falls". There is a "more" link below the list.

On the right side, there is a section titled "Movie and TV News" with the date "Thu February 6, 2003:" and a "Celebrity News" sub-section. The "Celebrity News" section includes links for "Catherine Deneuve Injured in Car Crash", "Jackson 'Devastated' by Documentary", and "Virgin and Love Kiss and Make Up". Below this is a "Studio Briefing" section with links for "Hollywood Baffled by War", "MGM Roars Over Strong Quarter", and "Hoffman Receives British Film Award".

The browser's status bar at the bottom shows "Document: Done (12.065 secs)".

Example 6

The screenshot shows the Wikipedia Main Page as it appeared in a Netscape browser window. The browser title is "Main Page - Wikipedia, the free encyclopedia - Netscape". The address bar shows the URL "http://en.wikipedia.org/wiki/Main_Page". The page content includes a globe logo with the text "WIKIPEDIA The Free Encyclopedia", a navigation menu with links like "Main Page", "Community portal", and "Recent changes", and a search box. The main content area features a "Main Page" section with a welcome message, a "Today's featured article" section about the Liberal Party of Utah, and an "In the news" section mentioning Hunter S. Thompson. The browser's status bar at the bottom shows "Done".

Main Page - Wikipedia, the free encyclopedia - Netscape

File Edit View Go Bookmarks Tools Window Help

W http://en.wikipedia.org/wiki/Main_Page

Create account / log in

article discussion view source history

WIKIPEDIA
The Free Encyclopedia

- Main Page
- Community portal
- Current events
- Recent changes
- Random page
- Help
- Donations

search

Go Search

toolbox

- What links here
- Related changes
- Special pages

Wikimedia needs your help in its US\$75,000 fund drive. See [our fundraising page](#) for details.
We are over 35% there thanks to all those who have donated so far! [Day 2 detail](#) and [report](#)

Main Page

From Wikipedia, the free encyclopedia.

Welcome to [Wikipedia](#), a [free-content](#) encyclopedia that [anyone can edit](#).

In this English version, started in 2001, we are currently working on [481388](#) articles.

[CULTURE](#) | [GEOGRAPHY](#) | [HISTORY](#) | [LIFE](#) | [MATHEMATICS](#) | [SCIENCE](#) | [SOCIETY](#) | [TECHNOLOGY](#)

[Browse Wikipedia](#) - [Article overviews](#) - [Alphabetical index](#) - [Other category schemes](#)

Today's featured article

 The [Liberal Party of Utah](#), along with the [People's Party](#), was a local [political party](#) that flourished in [Utah Territory](#) in the latter [19th century](#) before [Democrats](#) and [Republicans](#)

In the news

- [American author and journalist Hunter S. Thompson](#) commits suicide at the



Example 7

The screenshot shows a Netscape browser window titled "ACM Digital Library - Netscape". The address bar contains the URL: <http://portal.acm.org/dl.cfm?coll=portal&dl=ACM&CFID=38532722&CFTOKEN=>

The main content area features a purple banner with the "acm" logo and the word "PORTAL" in large white letters, with "SASLI" underneath. To the right of the banner are links for "Subscribe (Full Service)", "Register (Free, Limited Service)", and "Login". Below the banner is a search section with a radio button selected for "The ACM Digital Library" and another for "The Guide". A search input field and a "SEARCH" button are present.

THE ACM DIGITAL LIBRARY

Full text of every article ever published by ACM.

- ◆ [Using the ACM Digital Library](#)
- ◆ [Frequently Asked Questions \(FAQ's\)](#)

Recently loaded issues and proceedings:
(available in the DL within the past 2 weeks)

- IEEE/ACM Transactions on Networking (TON)
[Volume 12 Issue 6](#)
- Queue
[Volume 3 Issue 1](#)
- ACM SIGAPL APL Quote Quad
[Volume 34 Issue 1](#)
- ACM SIGAda Ada Letters
[Volume XXIV Issue 4](#)
- ACM SIGMOD Explorations Newsletter

◆ [Advanced Search](#)

◆ **Browse the Digital Library:**

- ◆ [Journals](#)
- ◆ [Magazines](#)
- ◆ [Transactions](#)
- ◆ [Proceedings](#)
- ◆ [Newsletters](#)
- ◆ [Publications by Affiliated Organizations](#)
- ◆ [Special Interest Groups \(SIGs\)](#)

Personalized Services: [Login required](#)

- ◆ [My Binders](#)
Save search results and queries. Share binders with colleagues and build bibliographies.
- ◆ [TOC Service](#)

Example 8



Example 9

The screenshot shows a Netscape browser window displaying the kalahari.net website. The browser's address bar shows the URL <http://www.kalahari.net/HOME/front.asp?>. The website header includes the kalahari.net logo, the main banner "The BIG Phat Clearout now on!", and user links for "Sign in", "my profile", "my basket", "my order", "contact us", "help", and "quicklinks". A "my shop!" icon is also present.

The navigation menu consists of buttons for HOME, BOOKS, MUSIC, DVD & VIDEO, GAMES, ELECTRONICS, HEALTH, CAMERAS & ACCESSORIES, VERIMARK, VOUCHERS, Computicket, and SEE MORE SHOPS. Below this, there are three promotional buttons: "NEW TO KALAHARI? CREATE A PROFILE!", "EARN EXTRA CASH JOIN OUR E-TRADERS!", and "TERMS & CONDITIONS".

A horizontal banner below the navigation menu states "FREE DELIVERY FOR DOMESTIC ORDERS OVER R300" with a car icon and a "CONDITIONS APP" link.

The main content area is divided into several sections:

- SEARCH:** A search bar with a dropdown menu set to "All products" and a search button. Below it, a link says "Need help searching? Try our Search Wizard".
- New to Kalahari? Create a profile here** and **Shopped with us before? - Sign in here**
- The BIG Phat Clearout Huge savings now on!** A large blue banner with the text "Fiction | Children | Lifestyle | Travel | General Non-Fiction | Afrikaans | 24hr delivery books on sale" and "Music | DVDs | Videos | Games | Verimark | Photographic & Electronics". Below the banner are three book covers.
- In my basket:** A section indicating "Your basket is empty" with a "TO CHECKOUT" button, and links for "My Address Book" and "ExpressPay Details".
- My wishlists:** A section showing a book titled "Dalene Matthee 1938 - 2005" with a cover image.
- 24 hour delivery:** A list of product categories with discounts: Books - 20% off, CDs - 15% off, DVDs - 10% off, Games - 10% off, and Cameras.
- Kodak CX7310:** A link to a product page.

The browser's status bar at the bottom shows various system icons and a taskbar.

Example 10

The screenshot shows the Amazon.com homepage in a Netscape browser window. The browser title is "Amazon.com--Earth's Biggest Selection - Netscape". The address bar shows the URL: <http://www.amazon.com/exec/obidos/subst/home/home.html/1>. The page features the Amazon logo, navigation links for "VIEW CART", "WISH LIST", "YOUR ACCOUNT", and "HELP". A promotional banner for "Amazon Exclusive!! Order a Segway now! It's only at Amazon" is visible. Below the navigation, there are buttons for "WELCOME", "YOUR STORE", "BOOKS", "APPAREL & ACCESSORIES", "ELECTRONICS", "TOYS & GAMES", "DVD", "SOFTWARE", and "SEE MORE STORES". A blue navigation bar contains "INTERNATIONAL", "TOP SELLERS", "TARGET", "TODAY'S DEALS", and "SELL YOUR STUFF". A large orange banner advertises "Super Saver Shipping FREE on orders over \$25" with a note that "Restrictions apply".

The main content area is divided into several sections:

- SEARCH:** A search box with a dropdown menu set to "All Products" and a "GO!" button.
- BROWSE:** A section titled "Featured Stores" with links for "Apparel & Accessories" and "Office Products".
- Personalized Recommendations:** A message: "Hello. Sign in to get [personalized recommendations](#). New customer? [Start here](#)."
- Black History Month Music:** A promotional section for "Save up to 30% on Black History Month Music". It features two album covers: "A Love Supreme: John Coltrane" and "When the Sun Goes Down: The Roots World Circuit". Both covers have a "30% off" badge. The text says: "This month, save up to 30% on classic jazz, R&B, blues, and more in our [Black History Month Sale](#)."
- Shopping for Your Sweetheart?:** A section with a red heart icon and text: "Choose Valentine's Guaranteed Shipping. [See details](#)."
- WHAT'S NEW:** A section with a "7" badge and text: "Already a customer? [Sign in](#) to see what's".

The browser's status bar at the bottom shows "Document: Done (34.6 secs)".

Example 11

The screenshot shows a Netscape browser window with the title "Networked Digital Library of Theses and Dissertations - Netscape". The address bar contains the URL "http://hercules.vtls.com/cgi-bin/ndltd/chameleon". The page header features the text "NETWORKED DIGITAL LIBRARY OF THESES AND DISSERTATIONS" and navigation links for "SEARCHES Scan | Keyword | Expert | English". Below the header, there are links for "New Session | Save Session | Cart | View History | Help". The main content area is titled "Scan Search" and includes instructions: "Enter your search terms.", "Select a search type.", "Choose a target database.", and "Click the Search button." A search form is present with an input field, a dropdown menu set to "Author", and a "Search" button. The target database is identified as "NDLTD Collection". A welcome message follows, describing the project's goal to increase the availability of student research. To the right of the text is a 3D cube graphic with a world map on its top face. At the bottom, there is a link to a press release and a note about browsing the catalog by author, subject, title, and library call number. The browser's status bar at the bottom shows "Document: Done (17.316 secs)".

Networked Digital Library of Theses and Dissertations - Netscape

http://hercules.vtls.com/cgi-bin/ndltd/chameleon

NETWORKED DIGITAL LIBRARY OF THESES AND DISSERTATIONS

SEARCHES Scan | Keyword | Expert | English

New Session | Save Session | Cart | View History | Help

Scan Search

Enter your search terms. Select a search type. Choose a target database. Click the **Search** button.

 Author **NDLTD Collection** Search

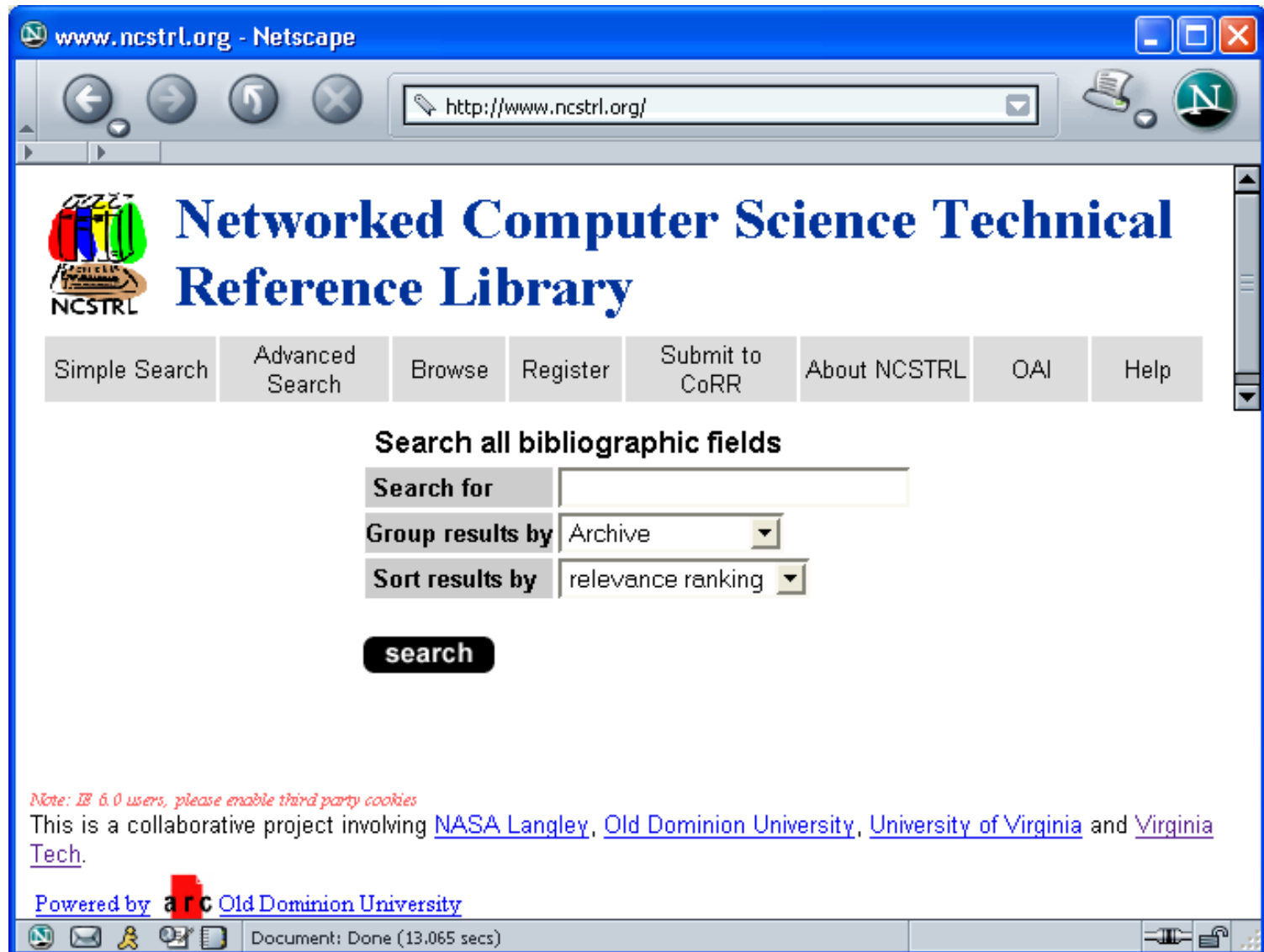
Welcome to the Networked Digital Library of Theses and Dissertations Union Catalog. This Union Catalog will serve as a repository of rich graduate educational material contributed by a number of member institutions worldwide. Our hope is that this project will increase the availability of student research for scholars, empower students to convey a richer message through the use of multimedia and hypermedia technologies and advance digital library technology worldwide.

This project is a joint development with [NDLTD](#) and [VTLS Inc.](#) Read the press release [here](#).

Browse the catalog by theses and dissertation author, subject, title and library call number.

Document: Done (17.316 secs)


Example 12



The screenshot shows a Netscape browser window with the address bar set to <http://www.ncstrl.org/>. The page title is "www.ncstrl.org - Netscape". The main content area features the NCSTRL logo on the left, which includes a colorful graphic of books and the text "NCSTRL". To the right of the logo is the heading "Networked Computer Science Technical Reference Library" in a large, blue, serif font. Below the heading is a horizontal menu with the following items: "Simple Search", "Advanced Search", "Browse", "Register", "Submit to CoRR", "About NCSTRL", "OAI", and "Help". Underneath the menu is a search section titled "Search all bibliographic fields". This section contains three input fields: "Search for" (an empty text box), "Group results by" (a dropdown menu currently set to "Archive"), and "Sort results by" (a dropdown menu currently set to "relevance ranking"). A black "search" button is positioned below these fields. At the bottom of the page, there is a note in red text: "Note: IE 6.0 users, please enable third party cookies". Below the note is a paragraph of text: "This is a collaborative project involving [NASA Langley](#), [Old Dominion University](#), [University of Virginia](#) and [Virginia Tech](#)." At the very bottom, it says "Powered by [arc](#) Old Dominion University". The browser's status bar at the bottom shows "Document: Done (13.065 secs)".

www.ncstrl.org - Netscape

http://www.ncstrl.org/

 **Networked Computer Science Technical Reference Library**

Simple Search Advanced Search Browse Register Submit to CoRR About NCSTRL OAI Help

Search all bibliographic fields

Search for

Group results by

Sort results by

search

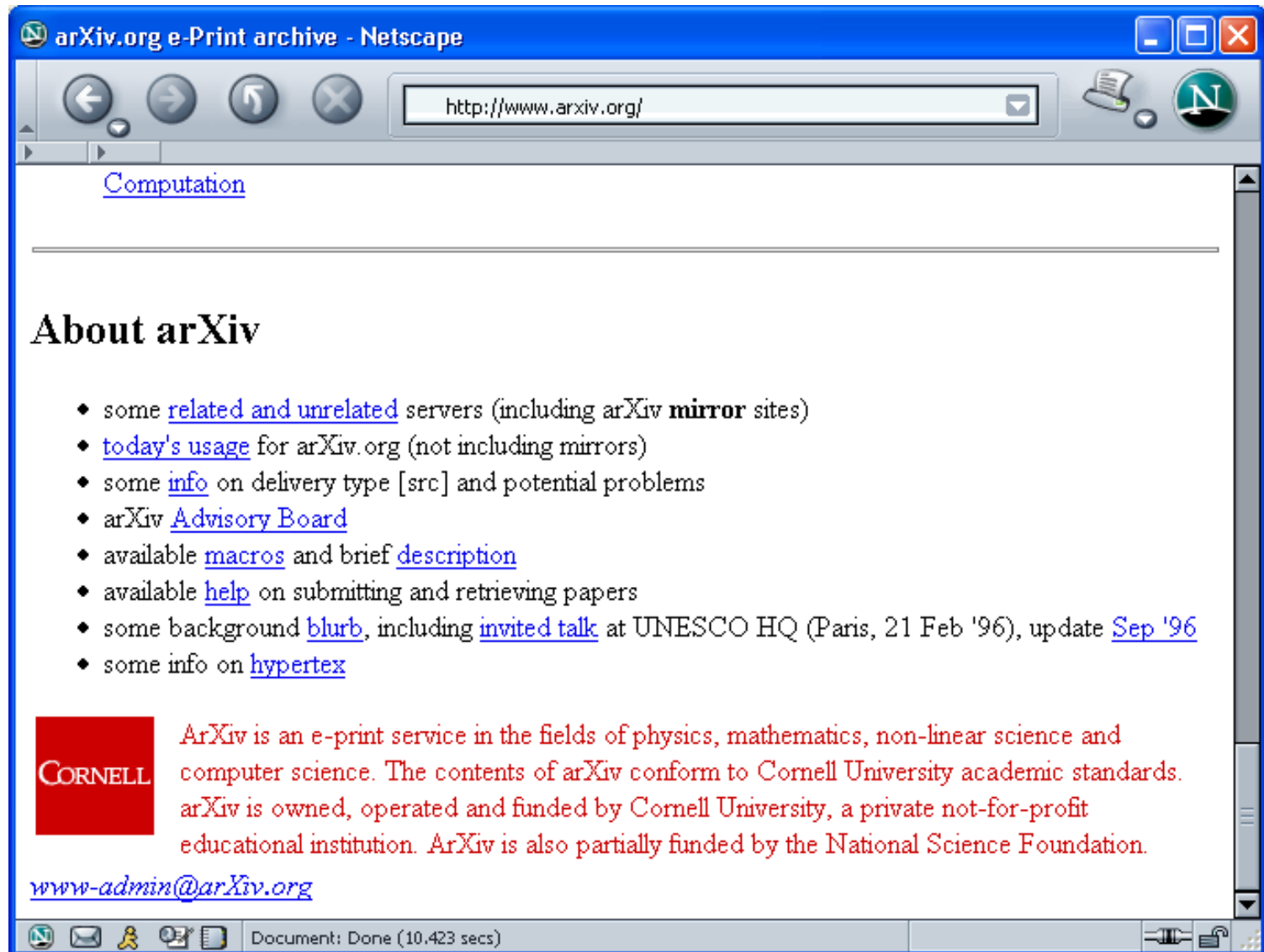
Note: IE 6.0 users, please enable third party cookies

This is a collaborative project involving [NASA Langley](#), [Old Dominion University](#), [University of Virginia](#) and [Virginia Tech](#).

Powered by [arc](#) Old Dominion University

Document: Done (13.065 secs)

Example 13




The screenshot shows a Netscape browser window titled "arXiv.org e-Print archive - Netscape". The address bar contains "http://www.arxiv.org/". The page content includes a link to "Computation", a section titled "About arXiv", a bulleted list of links, and a Cornell University logo with a paragraph of text. The status bar at the bottom indicates "Document: Done (10.423 secs)".

arXiv.org e-Print archive - Netscape

[Computation](#)

About arXiv

- some [related and unrelated](#) servers (including arXiv **mirror** sites)
- [today's usage](#) for arXiv.org (not including mirrors)
- some [info](#) on delivery type [src] and potential problems
- arXiv [Advisory Board](#)
- available [macros](#) and brief [description](#)
- available [help](#) on submitting and retrieving papers
- some background [blurb](#), including [invited talk](#) at UNESCO HQ (Paris, 21 Feb '96), update [Sep '96](#)
- some info on [hypertext](#)

 ArXiv is an e-print service in the fields of physics, mathematics, non-linear science and computer science. The contents of arXiv conform to Cornell University academic standards. arXiv is owned, operated and funded by Cornell University, a private not-for-profit educational institution. ArXiv is also partially funded by the National Science Foundation.

www-admin@arXiv.org

Document: Done (10.423 secs)

Example 14



Example 15

The screenshot shows a Netscape browser window titled "HowStuffWorks - Learn how Everything Works! - Netscape". The address bar contains "http://www.howstuffworks.com/". The website header features the HowStuffWorks logo (a stylized face with a globe for a head) and a navigation menu with categories: Computer Stuff, Auto Stuff, Electronics Stuff, Science Stuff, Home Stuff, Stuffo, Health Stuff, Money Stuff, and Travel. A search bar is located to the right of the logo.

Today @ HowStuffWorks

**Behind the Motion
How Electric Motors Work**

You might be surprised to find out just how much work is done by electric motors. They're everywhere these days! You'll find them in your car, your kitchen, bathroom, office, tool shed... If it moves, it probably has a motor.

[Find out](#) how electric motors work to produce motion.

Coming Tomorrow: How Spyware Works

How Kazaa Works
Should people have free access to movies, music and games? Kazaa has weathered the legal storm.

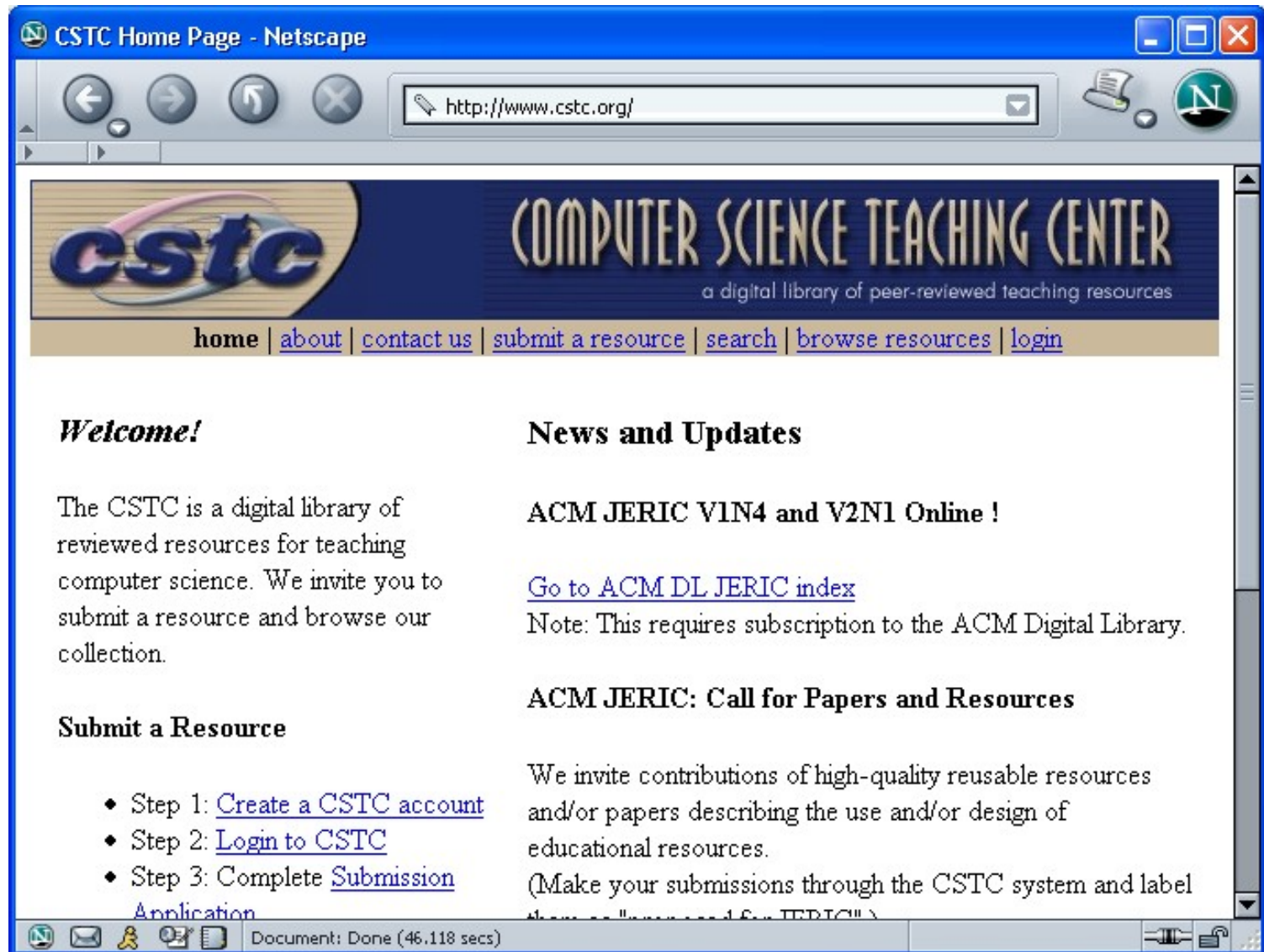
How the Ford SynUS Works
You'll be either giddy or horrified. City driving is bulletproof, comfortable and sporting a 45-inch LCD.

How Firewalls Work
A firewall is a critical component of a computer network. Wonder how it's protecting you from?

How the Jeep Hurricane Works
This concept off-roader has the hood and a turning radius that spins in place.

How Operating Systems Work

Example 16



The screenshot shows a Netscape browser window titled "CSTC Home Page - Netscape". The address bar contains "http://www.cstc.org/". The page features a header with the "cstc" logo and the text "COMPUTER SCIENCE TEACHING CENTER" and "a digital library of peer-reviewed teaching resources". Below the header is a navigation menu with links: [home](#), [about](#), [contact us](#), [submit a resource](#), [search](#), [browse resources](#), and [login](#).

Welcome!

The CSTC is a digital library of reviewed resources for teaching computer science. We invite you to submit a resource and browse our collection.

Submit a Resource

- Step 1: [Create a CSTC account](#)
- Step 2: [Login to CSTC](#)
- Step 3: Complete [Submission Application](#)

News and Updates

ACM JERIC VIN4 and V2N1 Online !

[Go to ACM DL JERIC index](#)

Note: This requires subscription to the ACM Digital Library.

ACM JERIC: Call for Papers and Resources

We invite contributions of high-quality reusable resources and/or papers describing the use and/or design of educational resources.

(Make your submissions through the CSTC system and label them as "submitted for JERIC")

The browser's status bar at the bottom shows "Document: Done (46.118 secs)".

Example 17

The screenshot shows a Netscape browser window displaying the NSDL website. The browser title is "NSDL - The National Science Digital Library - Netscape". The address bar shows "http://nsdl.org/". The website header includes the NSDL logo and a navigation menu with links: Home, Search, Collections, Of Interest, AskNSDL, About, Contact, and Community. A "Login" link is also present in the top right.

The main content area is divided into several sections:

- NATIONAL SCIENCE DIGITAL LIBRARY**: A large yellow banner at the top of the content area.
- NSDL PORTALS**: A blue sidebar on the left containing:
 - Middle School Portal**: Middle School Math, Science, and Technology presented by Eisenhower National Clearinghouse. Includes a link "What are NSDL Portals?".
 - FUTURE PORTALS**:
 - Teachers' Domain Pathways to Science**: Rich media for K-12 teaching presented by Teachers' Domain at WGBH.
 - The Math Gateway**: Undergraduate Mathematics presented by the Mathematical Association of America.
 - The Computational Science Education Reference Desk**.
- EXPLORE**: A yellow sidebar on the right containing:
 - AskNSDL**: How? What? Why?... Try the ask-an-expert service. Includes an icon of a head with a question mark.
 - New In The Library**: Holistic Numerical Methods. Includes the logo for the Holistic Numerical Methods Institute.
 - Resource Of Interest**: On the Cutting Edge Tsunami Visualizations. Includes a small image of a tsunami wave.
- HEADLINES**: A yellow banner at the bottom of the right sidebar.

A central image shows two young boys sitting at a desk with multiple computer monitors, looking at the screens. Below this image is the tagline: "The Nation's online library of resources for science, technology, engineering, and...".

The browser's status bar at the bottom shows "Done" and various system icons.

Example 18

The screenshot shows the Encyclopædia Britannica Online website as it appeared in a Netscape browser window. The browser's address bar displays the URL <http://www.britannica.com/>. The website's header includes the Britannica logo, the text "ENCYCLOPÆDIA Britannica Online", and navigation links for "Home", "Browse", "Newsletters", "Store", and "Subscri". A "Log in" link is also present for existing members.

The main content area is divided into several sections:

- What's New:** Features a "Presidential Quiz and Guide to Black History" and a notice about President's Day, encouraging users to "Test your knowledge of the U.S. presidents". It also highlights that February is Black History Month and provides a link to a "Guide to Black History".
- Search Britannica:** A central search box with a "GO" button. Below it, there are radio buttons for "Britannica Online" (selected) and "Dictionary & Thesaurus". Suggested searches include "baseball", "Greece", and "rocket".
- Other Ways to Find Information:** A grid of navigation options: "The Index", "Year in Review", "A-Z Browse", "Subject Browse", "World Atlas", "Timelines", "Classics", and "Dictionary Search" (with a Merriam-Webster logo).
- Biography of the Day:** Features a portrait of W.H. Auden and a "Learn More" link.
- This Day in History:** Features a historical image and text about the "Battle of Verdun begun" with a "Learn More" link.
- Britannica Highlights:** Features a portrait of George Washington and text about "Presidents' Day" honoring George Washington and Abraham Lincoln.

On the left side, there is a "Shopping" section advertising "New! The Encyclopædia Britannica World Atlas" for a price of USD \$99.95, described as a comprehensive and stunning world atlas.

The browser's status bar at the bottom shows various system icons and the taskbar.

Definition 1/5

- “Digital libraries are organizations that provide the resources, including the specialized staff, to select, structure, offer intellectual access to, interpret, distribute, preserve the integrity of, and ensure the persistence over time of collections of digital works so that they are readily and economically available for use by a defined community or set of communities.”
 - Greenstein, Dan (2000) DLF Draft strategy and business plan. Available <http://www.diglib.org/about/strategic.htm>

Definition 2/5

□ “Digital libraries are complex data/information/knowledge (hereafter information) systems that help: satisfy the information needs of users (societies), provide information services (scenarios), organize information in usable ways (structures), manage the location of information (spaces), and communicate information with users and their agents (streams).”

□ Fox, Edward A. (1999), DL Self-Study: definitions. Available <http://ei.cs.vt.edu/~dlib/def.htm>

Definition 3/5

- “Systems providing a community of users with coherent access to a large, organized repository of information and knowledge.”
 - Lynch, Clifford and Hector Garcia-Molina (1995), “Interoperability, scaling, and the digital libraries research agenda: A report on the May 18-19, 1995 IITA Digital Libraries Workshop”. Available <http://diglib.stanford.edu/diglib/pub/reports/iita-dlw/main.f>

- “The virtual or digital library is not an oxymoron- it is redundant. ... Since we did not bother to qualify our libraries by calling them clay libraries or papyrus scroll libraries, why now do we have to call them digital libraries?”
 - Braude, Robert and Samuel J. Wood (1999), “Virtual or actual: The term library is enough”, Bulletin of the Medical Library Association, p. 87.

Definition 4/5

- A digital library is “a world of literature, history, photographs, movies and maps open, free of charge, to any curious mind that wants to meander through the electronic equivalent of library stacks.”
 - Lipkin, Richard (1995), “The library that isn't there: Digital libraries transform books, photos, and videos into bits and bytes”, *Science News*, Vol. 147, No. 22, pp. 344-346.

Definition 5/5

- “a focused collection of digital objects, including text, video, and audio, along with methods for access and retrieval, and for selection, organization, and maintenance of the collection.”
 - Witten, Ian and David Bainbridge (2002), *How to Build a Digital Library*, Morgan Kaufman, p. 6.

So what is a Digital Library ?

- collections of digital objects
- palette of services
 - storage and preservation
 - access and use
- users
 - information seekers
 - information producers
 - information managers
- systems
 - network- and storage-based computer systems

Is the www a digital
library ?

Variety of Perspectives

□ Computer Science

- technical issues
- preference for automatic solutions e.g. Google

□ Library Science

- policies and organisational issues
- preference for human-mediated solutions, e.g. library cataloguing

□ Information Science

- philosophical issues ?

□ Physics, Chemistry, Medicine, Economics, etc.

- practical issues – how can we leverage digital libraries to solve our information management problems?

CSC400 Perspectives

- Digital Libraries are a conceptual space within which we define technology and policy to organise information effectively to address the needs of users.
- Digital Libraries provide a framework within which to devise advanced mechanisms for information management on the WWW and beyond.

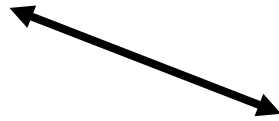
The Data and Services Model



DL as computer system

- ❑ Software package(s) to manage data (aka digital objects) and provide access to users, either locally or through a Web-based service.
- ❑ Software is used to
 - provide services to users
 - mediate data between layers
 - manage data storage and access
- ❑ Layers are not necessarily distinct!

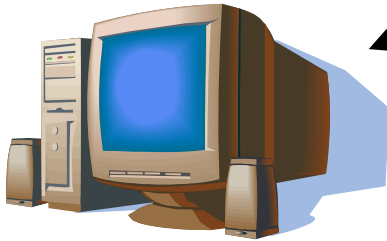
The 3-tier DL model



Services

Middleware ?

Data



Digital Object Types

Type	Example
Text	
Hypertext	
Image	
Video	
Audio	
3D Model	
Interactive Visualisation	
Software	

Common Features

- ❑ Can be created/destroyed
- ❑ Can be serialised and stored/retrieved electronically
- ❑ Can be transferred from one system to another (e.g.)
- ❑ Can be described
- ❑ Can be linked to (e.g.)

Examples of Services

- Google search
- Yahoo! directory
- Mailing lists
- Kalahari.net
- UCT Library catalogue
- -
- -
- -
- -

User management

- Authentication

- Check users are who they claim to be.

- Authorisation

- Check users are allowed to perform the tasks they are attempting.

- Maintain user information/profiles.

Searching

- Searching focuses on automatic/manual algorithms for indexing and querying.
- Indexing:
 - Transformation of information to support efficient discovery/retrieval.
- Querying:
 - Accessing transformed data to obtain results sorted in order of relevance, date, etc.
- a.k.a. Information Retrieval (IR)
- a.k.a. free-text databases
- Good example: Google
- Bad example: UCT website

Browsing

- Access subsets of data by categorical classification.
- Manual or automatic classification
- Single or multiple category membership
- Linear or hierarchical
- Is Searching = Browsing? Can searching be used as a surrogate for browsing?
- Example: Open Directory Project

Submission

- Add new digital objects to a DL.
- Content
 - digital objects
 - descriptions of objects
- Explicit submission vs. Harvesting vs. Crawling
 - Explicit submission = submission by local users
 - Harvesting = obtaining material from external sources
 - Crawling = finding material by automatically sifting through public collections e.g., WWW

Review

- Check submissions for appropriateness, quality, completeness, correctness, etc.
- Modes of review
 - Editorial review
 - Peer review
 - User review
- DL must support workflow for review processes.
- Security/Privacy issues must be addressed.
- Example: Online conference management

Annotation

- Add commentary or associated information to a digital object.
- Generalisation for reviews, ratings, discussions.
- May be stored as part of object or as separate objects.
- Link to objects and other annotations must be well-defined.
- Example: User feedback in online stores

Recommendation

- Suggest possibly relevant items based on past behaviour.
- Individual- or group-based recommendation
- a.k.a. Collaborative filtering (for groups)
- a.k.a. Selective Dissemination of Information (SDI) (for automatic push-services)
- Example: Amazon.com's recommended items

Middleware

- Databases and collections of data with standard/shared data formats
 - traditional approach
- APIs to access data, enabling use of different databases
 - current production systems
- Protocols to access data/services, enabling component-wise development of systems
 - current experimental systems

Why use 3-tier architecture ?
