WEB BROWSERS:
UNTANGLING THE WORLD WIDE WEB

Joan Latta Konecky
Carla Rosenquist-Buhler
University of Nebraska-Lincoln Libraries

Abstract

Not only are Internet resources expanding exponentially, but they are becoming more sophisticated, incorporating a variety of multimedia and hypertext components. Internet documents on the World Wide Web may contain elaborately formatted text, color graphics, audio, and video as well as dynamic connections to other Internet resources via hypertext links. In addition to providing user-friendly access to hypermedia resources, most Web browsers (client software) provide a rich graphical environment for authoring and displaying electronic documents locally.

This article describes the World Wide Web and a sampling of the available Web browsers. It then discusses a test project developed at the University of Nebraska-Lincoln Libraries designed to explore the potential, demands, and pitfalls of Web access to the Internet, as well as to investigate hypermedia document creation in an academic library environment. The experiences with the project confirmed the importance of the World Wide Web and Web browsers to this environment, so much so that providing access to these Internet resources must be seen as mandatory to any academic or upper level educational library providing electronic information access.

Introduction

In the rapidly evolving world of Internet, Mosaic gained enormous recognition in the popular media during 1993 as one of the brightest stars in Cyberspace. Due to its vivid appearance, enhanced capabilities and search friendliness, it made the World Wide Web's hypermedia capabilities readily accessible, quickly eclipsing the well known but simple Gopher software. In fact, media coverage of Mosaic led many to confuse the Mosaic software with the World Wide Web hypertext system that it searches. Mosaic is but one of many World Wide Web browsers developed in rapid succession to take advantage of the multimedia and hypertext nature of the Web. This article will describe both the World Wide Web and the client software designed to search it. It will also describe the Web Browser Demonstration Project at the University of Nebraska-Lincoln Libraries, which used the Mosaic client software to explore the capabilities and potential uses of World Wide Web browsing software in an academic library setting.

The World Wide Web

World Wide Web (aka WWW, the Web, or W3) is a unique Internet resource which uses hypertext to link documents on the network. It was developed in 1989 at CERN, the European Laboratory for Particle Physics in Switzerland, in order to distribute the literature of high energy physics electronically to researchers from many countries.1 "By creating computer linkage from the citations of an article to the corresponding source documents, users would be able to navigate through a body of related literature online simply by following the 'electronic footnotes', which would lead readers to the full text of the materials cited in those notes" 2

As the World Wide Web evolved, the hypertext links expanded beyond just footnotes in documents. Currently, hypertext links can appear anywhere within a document, allowing the user to make nonlinear jumps to related information anywhere, simply by "clicking" on a "link" -- either some highlighted text, a picture or an icon. The related information

Education Libraries • Vol 19, No. 1

13 —
may be located in the same document or may be in a source on a computer in a different country. To the user, the transition is transparent. Thus, the effortless hypertext links of the World Wide Web turn widely dispersed Internet sources into one vast seamless resource.3

In developing the World Wide Web, CERN employed three universal standards: HyperText Markup Language (HTML), Uniform Resource Locator (URL), and HyperText Transfer Protocol (HTTP). HTML was based upon the Standard Generalized Markup Language (SGML) and adds formatting tags to simple ASCII files. These formatting tags enable the document author to perform three functions: 1) to "display a variety of fonts, images, and highlighting options", 2) to indicate structural elements in the document, such as headers, lists, and paragraphs, and 3) to provide hypertext links to other documents on the Internet. 4

The Uniform Resource Locator (URL) is a unique identifier for each Internet resource and serves as the "address" for connecting to that resource. A URL has a standardized format of:  protocol://host/path/filename

The first part of the URL, protocol, describes the type of Internet resource, such as HTTP (WWW site or hypertext document), gopher, wais, telnet, ftp or file. The second part, host, is the Internet address and part of the computer where the information is stored. Path refers to the directory or subdirectory on that computer and filename refers to the specific file. It is very important to include the : and / punctuation in a URL.5

The underlying protocol used to transfer documents on the Web is HTTP. Unlike earlier Internet protocols, HTTP permits the transfer and retrieval of documents which have been formatted with Hypertext Markup Language (HTML) and which may contain graphic, audio or video elements.6 Computers that use HTTP to send or deliver HTML documents to users of the Internet are called "Web servers." Users browse the Web using "Webclient" software (such as Cello, Lynx, SlipKnot, Netscape or Mosaic) on their local computers.

The World Wide Web is a "client/server" system, composed of at least two separate computers, a "server" and a "client". A "server" is the central information provider which stores and, upon request from a "client", provides either the specific information resources or the hypertext links which point to information sources at other sites. One of the advantages of a client/server system, such as the World Wide Web, is that the user may choose from a variety of client programs, tailored for different needs, preferences and computing platforms.7

Accessing the Web: Web Browsers

The Web user interface (the way it looks, how it accesses the WWW site, and its ease of use) is determined by the Web client or "browser" software. There are many different types of Web browsers, tailored to the needs, or limitations, of the user and the user's equipment. The World Wide Web is a multi-protocol system that supports most of the earlier Internet protocols; therefore, searchers can use a Web browser to access and retrieve Internet resources via Gopher, FTP, WAIS, USENET, Archie, Telnet and others, without having to learn the intricacies of these other protocols.8 There are a growing number of World Wide Web browser packages, but among the more popular browsers are Lynx, Cello, SlipKnot, Mosaic and Netscape. These will be discussed briefly later.

Access to the World Wide Web became an overnight sensation when the graphical Windows-based Web browsers were developed, beginning with NCSA Mosaic, followed closely by Cornell's Cello, and later AirMosaic, Netscape, and a host of other competitors. These powerful graphical Web browsers make using Internet resources as easy as point and click.

With graphical Web browsers, users can access not only simple ASCII text but the full multimedia resources of the Internet, including formatted text, full-screen color graphics, audio recordings and video. These browsers use the WWW document "language," HTML, to display or print formatted text with headlines, subheadings, bold, italics, and different sizes and types of fonts. In most graphical browsers, the hypertext links are displayed in one color while previously selected links are remembered and displayed in another color.

Most Web browsers are designed with an open framework, using external viewer programs to handle the different types of textual and multimedia data. This open framework allows the browser software to 1) be customized with different individual viewer programs, 2) easily evolve to handle new data types, and 3) be able to integrate improved viewers. The Web browser software unifies these separate external programs, providing one consistent and easy to use interface for accessing many different types of Internet resources.9

Web browsers also provide many other capabilities. To improve the speed of access, many browsers store previously retrieved documents in cache memory on the local computer for rapid rebrowsing. An extremely valuable function of all the Web browsers is the ability to develop personalized Internet resource lists by saving the URLs to Bookmark or "hotlist" files for future reference. The URLs in the Bookmark lists are active hypertext links so that the user may jump directly from an entry in the Bookmark file to the specified Internet resource. These lists may also be exported, loaded or appended for others to use.

Network Access Issues

There are several issues which must be confronted when accessing the World Wide Web, including type of network connection, network traffic load and operating equipment requirements. First, there are several kinds of Internet connections from which to choose: dedicated or direct ethernet connection, SLIP or PPP connections and dial up "shell" (UNIX or terminal) connections.
Dedicated connections have their own IP address, use TCP/IP language, have unlimited access for multiple users, are very fast and expensive. SLIP (Serial Line Internet Protocol) or PPP (Point-to-Point Protocol) connections are available for individual dial up users and link directly to the Internet through the phone line. SLIP/PPP connections have complex installations which require the users to configure an IP address, load TCP/IP software, and, if they are using Windows, load Trumpet Winsock software. “Shell” account connections are simple, individual connections: the user connects to the UNIX host as a “dumb” terminal and uses the Internet software on the mainframe host, requiring little software or hardware investment. Typically, though, the “shell” account user can access only text-based Internet resources, such as FTP, email, and Gopher, and has access to the World Wide Web only through text-based browsers, such as Lynx and “WWW,” sacrificing the multimedia aspects of the Web documents. Recently though, the SlipKnot browser has circumvented the limitations of a “shell” account, providing full graphical, multimedia access to the Web.

Second, the combination of graphics, multimedia and text in the same document puts a heavy strain on the network traffic, producing a great deal of congestion at the most popular WWW sites. Transfer of multimedia and hypertext over the Internet requires a great deal of bandwidth and can be exquisitely slow, especially video. This increased load can put a tremendous burden on older network systems, producing system wide slow downs. It is barely tolerable even on a very powerful machine.

The third consideration for World Wide Web access are the equipment requirements. Generally speaking, though some graphical Web client software can be run on less powerful equipment, the recommended operating requirements are for a fast, well-equipped multimedia workstation:

**Windows version:**
- 80486 or Pentium processor
- 8MG RAM
- 200 MG hard drive
- Windows 3.2, and mouse

**MacIntosh version:**
- MacIntosh LCIII, System 7
- Color monitor

**UNIX version:**
- Any UNIX machine with true color support
- X-Windows

Despite these demands, the graphical browsers, like Mosaic, SlipKnot and Netscape, have become wildly popular. In addition, for commercial purposes, Netscape Corporation is offering technical support and more powerful server software which provide secure communications and server authentication for commercial transactions over the Internet. Many of the emerging Web browsers are designed to support business applications and reflect the change of focus in the Internet community from academic/research to home and commercial users.

A Sampling of Web Browsers

Web browsers are under rapid development. Within eight months, Mosaic, a browser that was hailed as a “killer application” in the mainstream media and popular journals, has been replaced by other critically acclaimed browsers. As new Web browsers emerge, new or improved capabilities are introduced in rapid fire, leapfrog fashion. Here is a overview of a few of the better known browsers.

**Lynx**
Lynx was developed at the University of Kansas in Spring 1993 and has DOS, VMS and UNIX versions available. It is a full-screen, terminal-based client. Because it is a non-graphical (non-Windows) browser, it allows access to the WWW from even the lowest common denominator VT100-based terminal or older, slower computer. Though Lynx can search and access the hypertext resources on the Web, it cannot access any of the graphic or multimedia components. Because it does not have to transfer the hypermedia content, it provides very rapid, efficient access to predominately textual Web resources.

**Mosaic**
The public domain Mosaic browser, created by the National Center for Supercomputing Applications, was first released as a UNIX X-windows client in early 1993. By Fall of 1993, Macintosh and Windows versions of Mosaic were available. This was the first graphical browser that supported hypertext access to the WWW information resources, while retaining support for the other Internet protocols such as Gopher, FTP, and WAIS. Originally, the Mosaic browser was problematic to install and configure, and was prone to crashing. The latest version, Enhanced NCSA Mosaic, is much more robust with improved capabilities and layout.

**Cello**
Cello is a public domain WWW client designed primarily for the electronic distribution of legal information. It was developed at The Legal Information Institute at Cornell University with a grant from the National Center for Automated Information Research in mid-1993. It requires a direct ethernet or SLIP/PPP connection to the Internet and only operates in the Microsoft Windows environment. Cello has many of the same features as the NCSA Mosaic client, such as History lists and Bookmark files, but also includes a straightforward print feature.

**Netscape**
The current successor to the Web browser throne of popularity is the Netscape Navigator browser. Netscape is a multiplatform product of Netscape Communications, a company formed by many of the original Mosaic development team. Netscape also requires a direct ethernet or SLIP/PPP connection to the Internet. The team reduced design flaws inherent in the
original NCSA Mosaic software and created a more efficient and powerful product. Netscape is visually more appealing, with improved menu structures, configuration options, and operational functionality. Most important, the Netscape browser performs better. Netscape appears to run more quickly than NCSA Mosaic, since the user can begin viewing a portion of the document before file transfer is complete. In addition, actual transfer speed was increased by improved screen layouts, automatic caching of previous pages, and streaming and compression techniques built into the Netscape browser client.¹⁵ Netscape was designed for a commercial clientele, though the browser is available free of charge to educational sites.

SlipKnot
SlipKnot is a shareware program that was developed by Peter Brooks of MicroMind, Inc., when he became frustrated by the trouble and expense of setting up a SLIP connection to the Internet. Unlike CellO, Mosaic and Netscape, the SlipKnot Web browser does not require a direct ethernet or complex SLIP/PPP (TCP/IP) connection to the Internet. SlipKnot operates under MS-Windows and provides full multimedia, hypertext access to the World Wide Web through a dial in “shell” (terminal or UNIX) account. SlipKnot uses Lynx or another text-based browser, “WWW”, to retrieve the Web documents and images. This “shell” or UNIX account access makes the World Wide Web resources readily available to a whole new group of Internet users, who only have indirect access to the Internet, such as dial in and home users. SlipKnot still requires a fairly powerful machine, at least a 386 with 8 megabytes of memory.

Though access is not as fast as a regular TCP/IP connection, SlipKnot can retrieve documents in the background while the user continues to work on other tasks. SlipKnot can display up to five documents at once and queues additional retrieval requests in the background while the user is viewing those documents. The retrieved documents may be saved on the local hard disk to be redisplayed at anytime, even offline.¹⁶ Although SlipKnot currently supports the FTP and HTTP (Web) protocols, it does not yet support Gopher, Mailto or HTML “Forms” capabilities, though this will probably be remedied in future versions.¹⁷

Demonstration Project
In the Fall of 1993, the University of Nebraska-Lincoln (UNL) Libraries brought up the Internet Resources Gopher, placing it as a menu option on the Libraries Electronic Information System/Online Public Access Catalog (OPAC). In addition to providing access to Internet information sources, the Libraries enhanced the Gopher by creating original cataloging entries for the Internet resources placed on the Gopher. In cooperation with the campus Computing Resources Center, the Libraries also linked the OPAC with the Campus Wide Information System (CWIS). This cooperative Gopher project promoted the Libraries as a knowledgeable provider of electronic information on campus, and the Libraries Gopher has proved to be popular and dynamic.

Gopher does not, however, support multimedia or hypertext. Since the UNL Gopher was introduced, a number of more powerful programs have emerged which take advantage of Internet hypertext and multimedia capabilities. A Windows Gopher client has been developed which supports graphics, but not hypertext. A number of World Wide Web clients which support hypertext and, in many cases, multimedia, have also been introduced in breathtakingly rapid succession (e.g. Lynx, CellO, Mosaic and Netscape).

Keeping pace with these rapid electronic developments and determining which are best suited to local needs are major issues for libraries with limited time, expertise and resources. In early 1994, the UNL Libraries’ Electronic Resources Program Group (ERPG) was charged with identifying, researching and making recommendations on current uses and future directions of multimedia technology in an academic library setting. In some ways this was similar to sketching an explosion, but, on the positive side, the project helped Library staff feel more comfortable and well-informed about developing technologies.

At that time, the World Wide Web was not well known, either by librarians or other faculty on campus, although the Web browser, Lynx—which supports hypertext but not multimedia—was available. Early on, the committee decided to focus their investigation on the capabilities and design flaws of Mosaic, which was the most user friendly and highly acclaimed Web browser then available.

In order to explore the capabilities of Mosaic, ERPG decided to create a “demonstration project.” They envisioned using Mosaic, with its expanded hypertext and multimedia capabilities, to provide services similar to those that Gopher was already providing, i.e. connecting to remote Internet sites, as well as the authoring and posting of library-related documents.

For this project, the ERPG Multimedia Demonstration Report was actually converted to a hypertext document using Hypertext Markup Language and made accessible on the Web. The report included links to documents on other Web servers to demonstrate remote searching and to show examples of sites that others were developing. The project document also contained several multimedia components, including graphics and a brief video. In this way, key issues related to both Internet access and local hypermedia authoring were tested.

The HTML version of the report is quite attractive, especially when compared to a simple ASCII gopher document. Graphics add visual interest and text is displayed in various font sizes and types. Bold facing provides emphasis and increases eye appeal and polish. The hypertext links allow users to access only the information they have an interest in, by “clicking” on those links.
Project Findings

The report, which was completed in the Summer of 1994, demonstrated that multimedia resources on the Internet were rapidly expanding, and recommended that the potential for more highly sophisticated academic uses justified continued investigations of Internet multimedia developments and World Wide Web use. However, the report also identified a number of key issues related to the implementation of Web access in the Libraries.

The first issues had to do with hardware necessary for implementation. Mosaic is resource intensive, needing fast, powerful computers and a high speed campus network. The network bandwidth required to transfer hypertext and/or multimedia is much greater than that of simple ASCII text via Gopher. This was a major problem identified campus-wide. Even the limited use of Mosaic by one of the high technology laboratories on campus produced a marked degradation of the campus network. In response to this immediate problem, as well as to prepare for projected network needs, the University embarked upon a complete re-wiring of the entire campus network with high speed Category 5 wiring. The Libraries were identified as one of the first sites to undergo rewiring.

Before a Web browser could be implemented, however, the Libraries had to upgrade many of the older, slower computers. This was already in process, due to the memory intensive requirements of many newer electronic resources. The Libraries are moving from a terminal based OPAC environment to a more flexible and powerful workstation environment for both public and staff use. Due to the expense of equipment upgrades, however, this will be an ongoing process over a number of years.

Another set of issues in fashioning a World Wide Web site for the Libraries revolved around the administration of the Web site, i.e. the establishment and maintenance of the system, creation of a Homepage or Homepages for the libraries on campus, development of Web structure and selection of sites. It was clear that the Libraries Automated Systems Office would assume primary responsibility for Web Administration, as it had for the Gopher, since this fell within the existing responsibilities. It was less clear who should be responsible for other duties. It did seem necessary, however, that a number of library staff would need to be involved with establishing the Web site; this included learning HTML, setting up homepages, developing library-related information screens, providing collection development and cataloging information and training other staff as well as the public.

Issues related to authoring Hypermedia documents were also considered. At the time of the demonstration project, there was little HTML expertise in the Libraries or on campus and no available documentation for converting a document into HTML. Only months later, several hypertext guides were posted on the Internet, somewhat easing this problem. The task of converting a linear document into a nonlinear hypertext format was also a conceptual challenge, requiring new ways of looking at document structure. How does one maintain the intellectual integrity of the document when the searcher has access to a multidimensional maze of interconnected information?

The complexity of capturing the multimedia components (audio, video, or graphics) in a digitized format required access to additional multimedia production equipment, computer support and training. Integrating these formats into a coherent and unified presentation was a real design challenge, for visualizing and constructing a multimedia guide is vastly different from preparing a linear print document. There were many questions about how various formats would coordinate and how they would appear on the screen.

Structuring information using hypertext requires a new way of looking at document organization. When developing the Gopher, the Libraries had applied existing collection development policies to the resources of the Internet. Subject specialists searched the Internet extensively to identify sites appropriate and useful to researchers. Gopher, through its two dimensional, hierarchical structure, provides inherent structure and control of Internet resources. Hypertext, however, provides multi-dimensional search possibilities. Could/should the Libraries attempt to structure and organize Web access in the same manner as the hierarchical Gopher? Given the range of materials available, should (or could) the Libraries attempt to restrict or "censor" access in any way?

In an attempt to improve access to Internet resources, the Libraries had developed an Internet Resources Gopher that included a locally-developed catalog of the Internet sites added to the Gopher. This catalog increased the access points to the Internet resources. The Libraries wondered if catalogers should also do this for Web sites? If so, should they continue cataloging Gopher sites as well? In fact, should the Libraries continue to maintain both a Gopher and a Web site, or should there be a plan to migrate to only a Web environment?

Providing public access to the World Wide Web has its own set of issues, including training, time and control. How active a role should librarians take in training the public? What would the training format be like (organized group sessions vs. on-demand individual sessions), and would the Libraries have the ability to provide training support for classroom assignments and in-depth research projects?

Hypermedia documents are necessarily large files which require a great deal of time to transfer. With some browsers, all parts of the document, text and graphics, must be imported before it will display. Even without a video component, this can be quite a lengthy wait. The speed of transfer is also impacted by the traffic clogging the network. Should the Libraries attempt to restrict the length of time a user can spend on a public access Web station if others are waiting? How many workstations will be needed to meet the public demand when sessions can be so lengthy? And a related issue, how many workstations can the Libraries "afford" to dedicate to Web browsing, without denigrating access to other electronic resources?
Dealing With Issues

The Mosaic demonstration project was extremely useful in giving the Libraries a greater understanding of the potentials and problems in using a hypermedia Web browser (either Mosaic or a similar product), both for providing expanded Internet access and as a document authoring tool. In order to deal with the issues and questions described above, the Libraries formed several committees to oversee various aspects of the introducing Web Browsers.

A Multimedia Implementation Committee was charged with making recommendations about which browser to implement, the placement of browsers, staff training needs, and training development. The Committee also developed Computer Use Guidelines to address issues such as time limitations and appropriate use. A Homepage Committee was created to begin developing the Libraries Web site. An existing Library Handouts Committee agreed to take on the task of creating informational documents using HTML.

The Multimedia Implementation Committee was still in the process of evaluating Web browsers when the first beta version of Netscape was released. The capabilities of even the early beta versions were so promising that the Computing Resources Center and the Libraries quickly selected Netscape rather than Mosaic as the browser of choice on campus. Currently, Netscape seems to be the most successful Web browser available. However, in the rapidly changing Internet environment, another browser could quickly take the lead.

As of this writing, the Libraries are in the process of establishing World Wide Web access on staff workstations. Due to current wiring, bandwidth and equipment capabilities, however, not all library staff will begin with Netscape access. Staff who do not need multimedia capabilities will continue to use Lynx, which functions well on existing equipment. Other staff members will use SlipKnot, which does not require a direct network connection and does not inhibit access to telecommunications software.

The Libraries Homepage is ready, though certain links are still under construction. The first public access Web stations, using the Netscape browser, should be available within weeks. The Libraries will start with a small number (probably 3-4) public browsers placed in the reference area of main library and one in each of the larger branches. As new workstations are purchased and the campus re-wiring is completed, additional public and staff Netscape workstations will be added.

Staff training is underway. All staff have had the opportunity to attend a general training session and, as they receive access to a Web browser, staff are given more in-depth training. In addition, the Libraries are in the process of building a new high-tech classroom. Recognizing the rapid proliferation of sophisticated electronic library research tools, as well as the increased training required to master these tools, the Libraries plan to use this room extensively to provide more comprehensive, hands-on computer training opportunities to both library staff and users. This room should be completed in the Fall 1995, and can be used for Web training as well as training on many other electronic resources.

However, plans have not been made yet for public training, rather, we are waiting to see what the need will be. The Libraries are adopting a flexible policy similar to the one that worked well in Gopher development; that is, NOT to attempt to have all the answers or anticipate all the related issues when embarking on the project. Although staff have tried to anticipate problems, in such a rapidly changing environment it is impossible (and even counter-productive) to plan for every eventuality. This is something of a paradigm shift for librarians who have in the past been expected to be "experts" (and who are accustomed to knowing the answers). A willingness to experiment and take risks are valuable attributes for librarians in this evolving electronic environment. It is likely that the librarians will be learning right along with (or even from) the users.

Conclusion

The Internet continues to expand in resources and in complexity. Tools like Gopher, which seemed powerful and innovative only a few years ago, are being superseded by next generation client/server programs (Web browsers), which access the hypermedia resources on the Internet. Simple ASCII documents are enhanced with a broad range of multimedia components and connected dynamically to other documents by hypertext links. Multimedia adds an additional layer of information to documents, augmenting their educational potential. But, perhaps more important, hypertext is redefining the way Internet users find and use information, changing the entire concept of what is meant by a "document" and eliminating traditional document boundaries.

The dynamic search potential of the World Wide Web browsers makes them critically important to scholarly pursuits. These browsers allow researchers a new freedom to relate and intellectually process pieces of information from multidisciplinary sources. Despite the challenges encountered in attempting to keep pace with increasingly sophisticated technologies, the University of Nebraska-Lincoln Libraries are convinced that providing access to these resources is vitally important to any academic or educational library which aims to stay at the forefront of scholarly information access.
Bibliography


To retrieve this article, send the following e-mail message to listserv@uhupvm1.uh.edu:
GET BARRY PRV5N5 F=MAIL
Or use the following URL:
gopher://info.lib.uh.edu:70/00/articles/e-journals/uhlibrary/pacsreview/v5/n5/barry.5n5


4. Ibid.


To retrieve this article, send the following e-mail message to listserv@uhupvm1.uh.edu:
GET MORGAN PRV5N6 F=MAIL
Or use the following URL:
gopher://info.lib.uh.edu:70/00/articles/e-journals/uhlibrary/pacsreview/v5/n6/morgan.5n6


Joan Latta Konecky is the Coordinator for Electronic Reference Services and a member of the Electronic Reference Program Group. Carla Rosenquist-Buhler is a Reference and Education Liaison Librarian and a member of the Multimedia Implementation Team. Both have been actively involved in Gopher and World Wide Web development at the University of Nebraska-Lincoln Libraries. The authors may be reached via e-mail: joank@unllib.unl.edu carlarb@unllib.unl.edu