6. SYMPOSIUM: MOSAIC AND THE INFORMATION SUPERHIGHWAY—A VIRTUAL TIGER IN YOUR TANK

Chaired by C. Michael Levy, *University of Florida*

Mosaic and the information superhighway: A virtual tiger in your tank

C. MICHAEL LEVY

University of Florida, Gainesville, Florida

The World-Wide Web (WWW) is a wide-area hypermedia information-retrieval initiative that aims to give access to a large universe of documents. Through software such as Mosaic, the WWW provides a simple, consistent means to access a variety of information stored as text, graphics, video, and sound files. This paper details the requirements for using this software, illustrates some current and planned hypermedia databases, and addresses issues that could affect the greatest possible utilization of the system.

Many years ago, a multinational oil company mounted an advertising campaign for gasoline, featuring a spirited tiger that represented the boost in power that people might expect if they used this brand of gasoline. We adopted this metaphor for our symposium because we sought to show how the present and anticipated features planned for the World-Wide Web (WWW) can empower users in much the same way—only here, we anticipate that the resulting benefits will be quite real.

Initially, I provide an overview of the WWW, a truly global database using Mosaic, a graphical public-domain interface. After placing its birth in context, I will describe important features of the system and explain what readers will need in order to become information consumers and information providers. Finally, I will discuss some issues now on the horizon that expand the value of the WWW, as well as others that may impede our fullest use of its capabilities.

Where Did it Come From?

The WWW is generally described as a wide-area hypermedia information-retrieval initiative that aims to give universal access to a large universe of documents. The WWW project's achievement in its short lifetime has been to give users of computer networks a simple and consistent way to access a variety of media. There are several different ways, using either text or graphics, to

access the WWW, the current favorites being graphical interfaces with exotic names such as Cello and Mosaic.

Access to the WWW is through the Internet. The WWW is related to the Internet, but they are not the same thing. The Internet is a network of networks, connecting thousands of smaller regional networks around the world. Although it is useful to think of the Internet as encompassing all of the machines, cables, and satellite links that connect them, it is useful to think of the WWW as referring to something more abstract—namely, the knowledge embodied in all of those machines.

Newcomers to networking are often surprised to learn that nobody owns, manages, or governs the entire Internet. The Internet has its roots in the cold war, when it was realized that a highly centralized system of information transfer would be a prime target for enemies who, with a couple of well-placed explosions, could prevent military and government leaders from doing their jobs. The RAND Corporation, a premier cold-war think-tank, proposed building a network that would have no central authority, and that could operate even if in fragments. From the outset, all nodes would be coequals in terms of sending and receiving messages. Messages would be divided into individually addressed electronic packets. Once a packet was sent, it would meander through the network in the general direction of its destination until it reached its recipient. If major chunks of the network were destroyed, the multiplicity of interconnections would enable messages to arrive intact eventually. In the big scheme of things, the near certainty that the information could actually reach its destination after passing through a rugged, but tattered,

Requests for reprints should be made to C. M. Levy, Department of Psychology, Box 112250, University of Florida, Gainesville, FL 32611-2250 (e-mail: levy@psych.ufl.edu).

system was more important than the speed advantages that a vulnerable, centralized system might provide.

These principles were tested in a small-scale way in 1968 by the National Physical Laboratory in England. A year later, the Advanced Research Projects Agency (ARPA), a unit of the U.S. Department of Defense, launched an extended test, which initially involved the connection of four supercomputers on the newly christened ARPANET. Because these supercomputers were expensive to purchase and maintain, it was quickly apparent that the scientific enterprise would benefit enormously if geographically dispersed researchers could access them remotely. The initial four nodes grew to 37 by 1972.

A curious thing happened as the network grew. The major traffic on the network changed from data streams destined for sophisticated number analysis into news, professional interchanges, and gossip. E-mail as we know it today was born, and mailing lists enabling messages to be widely broadcast quickly followed.

The system flourished because the communication protocols that emerged enabled behemoth computers from different manufacturers to speak a common language. Originally, this language was the ARPA's Network Protocol, but over time, the standard changed to TCP/IP, which involves two distinct entities. The Transmission Control Protocol (TCP) is responsible for organizing outgoing messages into streams of electronic packets and putting them back together in the proper sequence when they reach their destination. The Internet Protocol (IP) oversees the addressing and routing of the packets through multiple computer systems, including those networks using entirely different communication standards, such as Ethernet or X.25. At some point after the ARPANET was electronically linked with such other networks, the ensemble came to be known as the Internet.

The Internet flourished in the United States because the National Science Foundation (NSF) in 1984 became a major catalyst by laying down the infrastructure for its expanding usage. New users meant additional traffic on the network, and almost as soon as a major electronic extension was opened on what became known as the "NSF backbone," it filled to capacity. Five times between 1984 and 1990, the NSF contributed huge sums to provide larger, speedier linkages between major points on the North American continent. The NSF will not continue to contribute to the funding of this backbone after the year 2000, but the decision by Congress in 1991 to upgrade the infrastructure by nearly 5,000%¹ by 1996 may help to prevent the information highway from experiencing gridlock, perhaps into the next millennium.

The Internet also flourished because development of desktop (and, later, laptop) computers attracted legions of information-hungry aficionados and venture capitalists who sought to satisfy their hunger. A computer, a modem, and access to an existing Internet link were all that was needed to connect. Connections to the Internet cost the government sponsors nothing and the end users only a pittance because TCP/IP was in the public domain.

Using the Internet has always been a bargain. Unlike with the telephone system or commercial network providers, there are no charges to the end users for use of the Internet. Because there is nothing that remotely resembles an Internet Corporation or a U.S. Department of the Internet, "The Internet" cannot charge for anything. Everyone using the Internet is responsible only for buying and maintaining their own machine and their connection to the next link.

Although the Internet has had a presence for some time, the WWW is a relative newcomer. Tim Berners-Lee proposed it in 1989 to CERN (an acronym, expressed in French, for the European Particle Physics Laboratory cooperative) as a way in which to share research ideas and data throughout the organization, which was physically located in a number of countries. By the end of 1990, the first working code for the WWW was rolled out for a demonstration—and the rest, as they say, is history. By June 1993, for example, 100 sites had installed WWW servers and had become information providers, and this figure grew in the next year to 3,800 unique, publicly accessible sites.

Usage of the WWW has also grown at a phenomenal rate. The computer at the National Center for Supercomputing Applications (NCSA), where a great deal of WWW software is produced, received some 100,000 requests for information every week during mid-1993. By mid-1994, the figure regularly exceeded one million requests per week. By the end of 1994, it was anticipated that each month would bring at least two million new users to the WWW.

These impressive statistics seem to be the direct result of the WWW offering a very simple user interface for resources on the Internet that previously had been challenging to retrieve. This feature—in combination with the various public-domain graphical interfaces—has contributed to the enormous increase in network traffic that has occurred on the WWW ever since its earliest days.

What the WWW Can Do, and How it Works

The virtual tour of the White House, announced on television news worldwide, provides a generally useful demonstration of the WWW to information consumers. The tour starts with a display, shown in Figure 1. (Hardware permitting, these displays are rendered in full color on a computer monitor or printer.) In almost every WWW application, when the user's software (called a *client* or a *browser*) begins, a screen like this appears. The browser shown is Mosaic, which was developed at the NCSA and is widely used because it has been implemented for the major computing platforms (see Table 1).

An application may have many screens, each of which is often called a page. A page such as that illustrated in Figure 1 is known as a *home page*. It orients the user to everything else available in the application. This view of the White House home page shows certain words emphasized typographically (e.g., with boldface type or colored letters). These are *hypertext links*, areas on which

Figure 1. The White House Tour home page seen on a Windows Mosaic browser.

a user can click to direct the software to branch to another page that contains additional information. Hypertext links are not represented in any standard way. Some browsers indicate them as underscored characters, while others use special colors, and many enable the user to alter the default modes of representation.

Any WWW browser gives the user the power to branch to other pages using these hypertext links. The linked pages might be on the same host computer as the home page or they might be on multiple computers, each in a different country. The user doesn't need to know where the information is stored.

Mosaic and other graphical browsers offer the user far more capabilities than the hypertext links that text-only browsers provide. One can click on a small color photograph to enlarge it to a full-screen view, and then clip it, retouch it, adjust its contrast and hues, rotate it horizontally and vertically, store it for later use, or print it immediately. The quality of the print is determined primarily by the capabilities of the printer. Excellent renderings in black and white of 8.5×11 -in. photographs can be achieved on 600 dpi laser printers.

Mosaic also enables the user to access, play, and store digitized sound files. From the White House home page,

for example, a user with a sound card and speakers can listen to welcoming messages from the president and vice-president. The browser offers users the exciting opportunity to display moving images from films or video. Patience is virtue, since a 30-sec video can take 5 min or more to transmit. *Hypermedia* is the term coined to describe the capability of establishing linkages among text, graphics, sounds, and still and moving images.

Most browsers enable the user to send information as well as to retrieve it. Notice, for example, that the White House tour invites the user to sign the guest book. In other applications, this capability provides a way for the user to register, to voice concerns, or to order products; the user can, for example, put in interlibrary loan requests or purchase consumer goods.

What Can I Do With the WWW?

Tours of the White House, the dinosaur exhibit at the Honolulu Community College museum, current weather maps from the National Hurricane Center, and radiographic images from the University of Iowa Virtual Hospital are all fascinating sites. After their initial appeal wanes, psychologists will search for other sites to justify their use of the WWW. Krantz (1995) shows how he has

Table 1
WWW Text and Graphical Browsers

Type of Browser	Platform
Text-only	Dumb terminals connected to nearly any UNIX platform Smart terminals (e.g., IBM 3270) DosLynx for DOS environments Macintosh SEs and above with System 7x. Text-only using curses (e.g., VT100 or IBM 3270 emulators) for SUNOS 4.1.x, IBM AIX, DEC OSF/1, DEC Ultrix, and VAX Multinet Browsers written for the emacs environment and written in perl are also available
Graphical	Mosaic, Netscape, and Cello for IBM compatibles (80386+ with 4-mb RAM, under Windows 3.1 in enhanced mode) Mosaic and Samba for Macintosh System 7.x, 68020+, and Power Macintosh OmniWebb (for NeXT, NeXTStep 3.0) Midas, Viola, and Chimera for X/DecWindows Mosaic for UNIX-based machines, high performance Sun, Silicon Graphics, and DEC workstations

provided a way to access a great deal of professionally relevant information that will also interest graduate and undergraduate students. This information ranges from funding and employment opportunities to on-line instruction and evaluation. The WWW offers unique special features for those who will primarily be information consumers, as well as to those who will also become information providers.

As Chu, Walter, and Palya (1995) show, organizing and formatting information for delivery systems requires some basic understanding of screen design, as well as the will and commitment to mark up the source copy, using the special *html* language needed to create hypermedia links and to format text. Fortunately, tools are now available for automating the markup process. For example, Microsoft has recently announced an extension to Word for Windows 6.0 that will accomplish this task. Apart from the substantial job of preparing documents for distribution via the WWW, there are a host of logistical and legal matters that need to be considered. Allan and Kostenbader (1995) address some of these. Because the same machine can be used to service local network needs and to provide data to the rest of the world, Allan and Kostenbader also discuss several possible strategies for arranging information on WWW servers.

The research implications of the capability to make our current work available electronically are impressive. For example, once I have established an electronic mailing list of colleagues who follow my work, it is a simple matter to write one note that I can send by e-mail to notify everyone that I have completed a new paper. Then, whenever any of them has time, they can use the WWW to read my report—complete with its colored graphics, figures, and formatted tables—and save it on their hard disk or print out a copy if they choose. Since all of these network operations are currently without cost to most ac-

ademic users²—either on the information-providing or the information-consuming end—department chairs and deans are thrilled because their postage costs drop significantly. The electronic services provided by the American Psychological Society, the American Psychological Association, and various Federal agencies through the Internet can help us to search for appropriate new funding sources and keep abreast of application deadlines.

The WWW may grow to have an important role in our teaching missions. Dale Spender (1994), the feminist writer, recently described the experiences of students at a small college for women in Australia who had networked laptops at their disposal. These experiences radically altered how the students approached their learning and transformed how the faculty approached its teaching. For example, students decompartmentalized their learning, practicing their French by turning in their math assignments in French. Fixed class schedules were abandoned, as well as fixed rooms for traditional instruction. Students spontaneously congregated in various unused indoor and outdoor spaces, and instructors moved between them to provide help and guidance when requested. Closer to home, my own university recently established a writing center that will be used in all beginning English composition courses. Every student will have access during and after class to networked workstations that will help them not only in searching electronic databases for citations and abstracts, but also in retrieving entire documents and still images, movie clips, and sound bites that they can edit and import when they compose hypermedia “papers.”

There is considerable freedom in the layout and formatting of hypermedia documents on the WWW, as is appropriate in a system that, in part, flourished because of its anarchistic roots. But standards and guidelines may soon be forthcoming. The Library of Congress has embarked on a program to digitize several million documents in its holdings each year and to make them available via the WWW (Lewis, 1994). Great thought will certainly go into the formulation of internal standards that may well become public guidelines. This may be as valuable to the information provider, as the stylistic conventions published in the *Publication Manual of the APA* are to authors who produce printed documents by making the *process* of communication a priority.

There are also administrative uses that psychologists in academia and in the private sector may recognize as beneficial. All manner of documents—catalogs, course schedules, telephone directories, events calendars, maps, and so forth—are being placed on servers that can be accessed locally with the same software used to access information remotely. Over time, it should be possible to reduce the number of copies of these documents that are printed. Eventually, some may be eliminated altogether in printed form.

Having detailed information about the institution and its researchers available through the WWW should enable institutions to recruit more effectively. Students and applicants will be able to learn about researchers’ cur-

rent projects, rather than having to rely on searches of databases for pointers to previously published work. This may help them to narrow their searches more intelligently, which may help to enhance the quality of the pool of applicants. The same function would also likely be served when a unit seeks to fill faculty and administrative vacancies, because with better-informed applicants, interviews can focus more on their professional promise than on what resources the unit has to offer.

What Do I Need in Order to Sign On?

As Table 1 indicates, there are browsers for the WWW for almost every hardware environment. The critical element, however, concerns the need to be able to connect to another computer that is itself linked to the WWW. The best response time will be achieved through an optic-fiber connection to a network host, but ordinary telephone lines and a modem can provide basic service.

Software options include text browsers (e.g., Dos-Lynx) and graphical browsers (e.g., Mosaic, Netscape, and Cello). This article and the three companion papers by Krantz (1995), Allan and Kostenbader (1995), and Chu et al. (1995) focus on the popular Mosaic software. Mosaic is a significant tool not only because of its hypermedia linkages, but also because one can use it to access many otherwise incompatible network services, such as Gopher, WAIS, and ftp (the file transfer protocol). For-profit software publishers are now beginning to offer their own competing products. One early player is Microsoft, which recently announced an add-on module for Word for Windows that will enable that product to browse the WWW. Other commercial products are in various stages of production.

Readers who plan to connect to their host through ordinary telephone lines will need to ensure that their host can run software such as SLIP (serial line Internet protocol) or PPP (point-to-point protocol) for access using a dial-in connection via a modem. To use a text browser, almost any modem will do, but if you want to use a graphical interface and be able to see still pictures or movies, any modem slower than 14K baud will require the patience of Job.

Local computer centers or MIS departments are the first places to search for help in getting started. They may hand you a disk that you only need to insert into your machine to have all of the necessary software installed and quickly readied for use.

What Do I Need in Order to Become an Information Provider?

The minimum requirements to become an information provider are not insurmountable for most colleges and departments: a reasonably fast desktop PC running almost any version of UNIX, Windows 3.1, or Windows NT, or a Macintosh or Power Macintosh running System 7.x, 68020+. A high-end Macintosh or 486-class PC can handle several hundred requests for information an hour without severely degrading word processing that occurs simultaneously. Appropriate server software is

also needed, and it is available free from the sources listed in the next section.

Getting More Information About the WWW

Without having a browser, the best way to get general information about the WWW is to access CERN, the organization that conceived it. The simplest ways to do this are to use telnet or gopher by typing either telnet info.cern.ch or gopher info.cern.ch. Information on how to obtain software is also found there. Readers unfamiliar with the use of ftp, gopher, or telnet can get a copy of the *Unofficial Internet Book List* simply by sending a request for it by e-mail to savetz@rahul.net.

Some Final Thoughts

I have three broad concerns about the WWW that need addressing. They concern (1) the future costs of access, (2) the changing role of publishing, and (3) the difficulty that new users often have in finding information.

1. I stated earlier that use of the Internet is generally free for individual academic patrons. For those without such access, commercial service providers such as America OnLine, CompuServe, and Prodigy are available for modest fees. However, the concept of free access to academic end users is currently being questioned, and the studies being conducted in response to those questions may have implications that reach far beyond academic institutions. Until fairly recently, the bulk of the cost of running the Internet in the U.S. was borne by the NSF, with universities contributing a smaller portion. The expanded National Information Infrastructure is presently developing along a rocky road. The basic issue is that it will be built only with the participation of the telephone and cable carriers, and these organizations will not participate if they have to compete with the government. Accordingly, by mid-1995, the NSF will abandon its role in the management of the network, and five years after that, will discontinue all funding. As the network becomes privatized, we have to consider where the revenue will come from to pay for our usage.³ A model currently favored by the private sector is one that involves billing for network usage on the basis of very small message units. The critical term is *usage*. In many quarters, this is interpreted to mean that when information travels across the network, both the information consumer and the information provider could be charged. This could kill access to the network as we now know it. There would be powerful disincentives to being an information provider to users outside the local organization if this model were adopted. If providers had to pay every time someone accessed their hard disk remotely, many information servers would simply be shut down.

Even if the user, rather than the information provider, were billed for the network access, and even if the charges were minimal, the mechanism for billing and collecting for a few cents here and a few cents there is

presently not economically feasible. But computer scientists at Carnegie Mellon University have publicly described a means whereby it can be profitable to bill a user's credit card for such very small amounts, and several commercial network providers have expressed interest in this approach. As costs are spread across an increasing number of users, we may see no increment in our local billing, but this is a volatile situation that warrants close monitoring.

2. The WWW will force a self-examination by the publishing industry, which may change the way it does business. The entry of the Library of Congress (LOC) into the field as an electronic information provider will virtually require such a self-examination. The LOC has vast holdings, but it will stay away as long as possible from knotty problems that involve copyright (Lewis, 1994). Some journals owned by professional societies will move with the times and simultaneously publish print and electronic forms of each issue. *The Florida Entomologist*—a leading international journal—has pioneered such dual publications using an inexpensive proprietary product, Acrobat2. The software reader needed by the subscriber to download, view, and print entire issues is available gratis. This journal has demonstrated the ease with which high-resolution electronic microscopic graphics can be transmitted and quickly printed on ordinary (i.e., non-Postscript) laser printers. Subscribers no longer need to request reprints; they can produce them to order on their own desktops.

3. My final concern is with the difficulty that newcomers to the WWW experience in locating the information that they know is there somewhere. Part of their problem is the result of there being no analog of a card catalog of all of the information available. Currently, there are WWW "Yellow Pages," but these are generally incomplete, even for small knowledge domains. Recall that no one is in charge of the Internet or the WWW, so at some level, this incompleteness is perfectly understandable.

Fortunately, some innovative people have designed useful tools that can help. The home page of my college at the University of Florida has a robot program that travels the WWW daily to identify colleges, universities, and federal agencies that have just become information providers. As a result, anyone who accesses this page (address: <http://www.clas.ufl.edu>) can quickly branch to the home page of any institution that the robot has discovered.

Another, more generally useful, tool is the World Wide Web Worm (WWWW; accessed at <http://www.cs.colorado.edu/home/mcbryan/wwww.html>). This is a robot capable of locating objects (text, pictures, sounds, or movies) that meet specific criteria. Perfectly reasonable requests to the WWW would be to identify all of the sound clips associated with home pages of psychology departments or to locate all of the movie clips associated with home pages in France.

When this article is published, the WWW will be only 4½ years old. It will have more than 2 million information providers and 50 million users. If a nation or an economy or a business were to experience such an explosive growth rate, there would probably be widespread concern. Instead, in this exceptional instance, there seems to be widespread enthusiasm.

REFERENCES

- ALLAN, R. W., & KOSTENBADER, P. S. (1995). Information on the Internet: How selective should we be? *Behavior Research Methods, Instruments, & Computers*, *27*, 198-199.
- CHU, J. Y. M., PALYA, W. L., & WALTER, D. E. (1995). Creating a hypertext markup language document for an information server. *Behavior Research Methods, Instruments, & Computers*, *27*, 200-205.
- KRANTZ, J. H. (1995). Linked Gopher and World-Wide Web services for the American Psychological Society and Hanover College Psychology Department. *Behavior Research Methods, Instruments, & Computers*, *27*, 193-197.
- LEWIS, P. H. (1994, September 12). Library of Congress offering to feed data superhighway. *New York Times*.
- ROBERTS, M. M. (1994). Why is the Internet so cheap? *Educom Review*, *26*, 6-8.
- SPENDER, D. (1994, May). *Ladies and laptops*. Plenary speech presented at the Tenth Annual Computers and Writing Conference, Columbia, MO.

NOTES

1. Assuming only limited traffic on the Internet, this backbone capacity will permit the transfer of the entire contents of the *Encyclopaedia Britannica* in one second.
2. Many universities hide the true cost of connecting to the Internet, paying for it with funds off the tops of their budgets. The average cost per university for Internet access averages less than a dollar per month for each student, faculty, and staff member (Roberts, 1994).
3. Unlike the case with a real highway, mere usage of the so-called information superhighway will not wear it out. The costs that we must bear represent our share of the support of all those satellites, optic fibers, and telephone and cable equipment, as well as the human resources required to install and maintain them.

(Manuscript received December 14, 1994;
revision accepted for publication January 9, 1995.)